Co-inheritance of sea age at maturity and iteroparity in the Atlantic salmon vgll3 genomic region

Aykanat, Tutku

2019-04

Aykanat , T , Ozerov , M , Vähä , J-P , Orell , P , Niemelä , E , Erkinaro , J & Primmer , C R 2019 , ' Co-inheritance of sea age at maturity and iteroparity in the Atlantic salmon vgll3 genomic region ' , Journal of Evolutionary Biology , vol. 32 , no. 4 , pp. 343-355 . https://doi.org/10.1111/jeb.13418

http://hdl.handle.net/10138/324094
https://doi.org/10.1111/jeb.13418

acceptedVersion

Co-inheritance of sea age at maturity and iteroparity in the Atlantic salmon *vgll3* genomic region.

Tutku Aykanat[1,2], Mikhail Ozerov[2,3], Juha-Pekka Vähä[3,4], Panu Orell[5], Eero Niemelä[5], Jaakko Erkinaro[5], Craig R. Primmer [1,6,7]

[1] *Organismal and Evolutionary Biology Research Programme, PO Box 56, 00014 University of Helsinki, Finland.*

[2] *Department of Biology, University of Turku, Turku FI-20014, Finland.*

[3] *Kevo Subarctic Research Institute, University of Turku, Turku FI-20014, Finland.*

[4] *Association for Water and Environment of Western Uusimaa, POB 51, FI-08101 Lohja, Finland.*

[5] *Natural Resources Institute Finland (Luke), PO Box 413, FI-90014 Oulu, Finland.*

[6] *lnstitute of Biotechnology, University of Helsinki, Finland.*

[7] *Helsinki Institute of Sustainability Science, University of Helsinki, Finland*

Short running title: Genetic correlation between life history traits

Corresponding author: Tutku Aykanat

Phone: +358 46 8962176

E-mail: tutku.aykanat@helsinki.fi

Acknowledgements

We acknowledge the fishers of the Teno River who contributed scales and phenotypic information to the Natural Resources Institute Finland (LUKE). Scale analyses were carried out by Jari Haantie. Jorma Kuusela is thanked for help with scale selection, analysis. Meri Lindqvist, Heli Junes, Jan Gerwin, Minna Kuusela and Kristiina Haapanen are thanked for laboratory assistance. This study was supported by grants from the Maj and Tor Nessling Foundation (project number 201600445 to TA), and Finnish Academy (grant number 318939 to TA, 286334 to JE and, 284941, 307593, 302873 to CRP). Victoria L. Pritchard is thanked for analysis suggestions. Data and codes related to model optimization and parameters are available in the Dryad Digital Repository: XXX-XXX.

Co-inheritance of sea age at maturity and iteroparity in the Atlantic salmon *vgll3* genomic region.

Abstract

Co-inheritance in life history traits may result in unpredictable evolutionary trajectories if not accounted for in life-history models. Iteroparity (the reproductive strategy of reproducing more than once) in Atlantic salmon (*Salmo salar*) is a fitness trait with substantial variation within and among populations. In the Teno River in northern Europe, iteroparous individuals constitute an important component of many populations and have experienced a sharp increase in abundance in the last 20 years, partly overlapping with a general decrease in age structure. The physiological basis of iteroparity bears similarities to that of age at first maturity, another life history trait with substantial fitness effects in salmon. Sea age at maturity in Atlantic salmon is controlled by a major locus around the *vgll3* gene, and we used this opportunity demonstrate that these two traits are co-inherited around this genome region. The odds ratio of survival until second reproduction was up to 2.4 (1.8-3.5 90% CI) times higher for fish with the early-maturing *vgll3* genotype (*EE*) compared to fish with the late-maturing genotype *(LL*). The *L* allele was dominant in individuals remaining only one year at sea before maturation, but the dominance was reversed, with the *E* allele being dominant in individuals maturing after two or more years at sea. *Post hoc* analysis indicated that iteroparous fish with the *EE* genotype had accelerated growth prior to first reproduction compared to first-time spawners, across all age groups, while this effect was not detected in fish with the *LL* genotype. These results broaden the functional link around the *vgll3* genome region and help us understand constraints in the evolution of life history variation in salmon. Our results further highlight the need to account for genetic correlations between fitness traits when predicting demographic changes in changing environments.

Keywords: C*o-inheritance, life-history evolution, Atlantic salmon, iteroparity, sea age at first maturity.*

Introduction

Since being formally described, multivariate evolution has been well incorporated into quantitative genetic frameworks, covering predictions under diverse theoretical scenarios (Lande, 1979; Wagner, 1989; Houle, 1991; Roff, 1996; Griswold & Whitlock, 2003; Chevin *et al.*, 2010; Wang *et al.*, 2010). Accordingly, the theory suggests genetic correlation between traits can constrain the pace and efficacy of natural selection (Lande & Arnold, 1983; Roff, 1996; Orr, 2000). When correlated characters have contrasting fitness trajectories in the adaptive landscape, the climb towards the local fitness peak can be restricted, and result in suboptimal fitness of populations (Lande, 1982). The principles of the multivariate theory of evolution have been successfully applied to many fields, such as animal and plant breeding, multi-trait artificial selection (Kadarmideen *et al.*, 2003; Careau *et al.*, 2010; Chen *et al.*, 2011; Kause *et al.*, 2011; Weigel *et al.*, 2017), and epidemiology (Lee *et al.*, 2012; Sanchez-Guillen *et al.*, 2012; Bulik-Sullivan *et al.*, 2015; Gratten & Visscher, 2016; Schnurr *et al.*, 2016; Hammerschlag *et al.*, 2017).

Measuring multi-trait evolution in wild populations has substantial importance. The magnitude and sign of genetic covariation between fitness traits may facilitate or constrain adaptive evolution by conflicting trait co-evolution, or so-called trade-offs (Roff, 1996; Sheldon *et al.*, 2003; Hellmann & Pineda-Krch, 2007; Agrawal & Stinchcombe, 2009; Duputie *et al.*, 2012; Chirgwin *et al.*, 2015). A handful of studies have investigated the genetic relationships between fitness-related traits in the wild (Sheldon *et al.*, 2003; Theriault *et al.*, 2007; Carlson & Seamons, 2008; Nussey *et al.*, 2008; Robinson *et al.*, 2009; Clements *et al.*, 2011; Lane *et al.*, 2011; Santure *et al.*, 2013), most of which were confined to well-studied populations in isolated or historically monitored settings with a scope to demonstrate the evolution of antagonistic genetic correlation between fitness traits (Roff, 1996). Indeed, accurately estimating genetic relatedness in the wild is challenging and requires either well-established pedigrees (Kruuk *et al.*, 2000) or the genotyping of large numbers of genetic markers in large datasets (Ritland, 1996; Yang *et al.*, 2010; Robinson *et al.*, 2013; Berenos *et al.*, 2014). Thus,

examples from wild populations are limited. Furthermore, multivariate genetic models are much more data-demanding than univariate models, making estimating genetic correlations more difficult than measuring univariate additive genetic variation (Roff, 1996; Nussey *et al.*, 2008; Wilson *et al.*, 2010).

Despite these challenges, measuring multivariate trait evolution in the wild has substantial implications for conservation and management efforts, e.g., to better predict population responses to changing environmental conditions (Etterson & Shaw, 2001; Hellmann & Pineda-Krch, 2007; Chirgwin *et al.*, 2015). In general, if a significant predictor of population demographic structure is left unaccounted, predictions may be inaccurate (Walters & Maguire, 1996; Dunlop *et al.*, 2009). In this sense, any potential genetic correlation between fitness-related traits will mediate the evolutionary response of populations and alter the trajectory and the pace of evolution towards the fitness optimum, perhaps in an unexpected direction. Also important, but often overlooked, in wild populations is the fact that co-inheritance between fitness traits may help to uncover the physiological basis of the correlation at the molecular level (Stearns *et al.*, 1991; Storz *et al.*, 2015). This can help illuminate the response of populations to ecosystem level processes and facilitate understanding how ecological dynamics shape trait evolution (Arnold, 1981; DeLong, 2017; Stearns, 2010; Scott *et al.*, 2015), see also (Storz *et al.*, 2015). In humans, for example, a recent GWAS study showed insomnia, and psychiatric and metabolic traits are co-inherited at a locus, suggesting that these traits are functionally linked by genetic determination (Hammerschlag *et al.*, 2017). Another study, again in humans, demonstrated pleiotropy between early- and late-life diseases, suggesting common etiology of these diseases (Rodriguez *et al.*, 2017). As such, these studies demonstrate pleiotropy as a powerful tool for guiding functional inference of genes, but such approaches are yet to be employed in the wild non-model systems.

Indeed, a significant source of genetic covariance between traits can arise from the co-inheritance of traits due to pleiotropic gene action or strong co-localization of causal genetic elements due to their

close physical proximity in the genome. Despite the two mechanism having different causations, marker assisted ecological studies use the term pleiotropy to describe both processes inclusively (e.g. Albert *et al.*, 2008). In line with this presupposition, true pleiotropy (trait co-variation caused by the same mutational element) and tight-linkage have similar co-inheritance patterns and evolutionary constraints, but potentially with different rates of (e.g. Paaby & Rockman, 2013).

Sea age at first maturity and iteroparity are two life history traits with profound effects on both reproductive output and survival in Atlantic salmon (*Salmo salar*, Fleming, 1996; Fleming & Einum, 2010) and other salmonid fish species (Christie *et al.*, 2018). Both exhibit substantial variation, likely as a result of fitness trade-offs maintaining the variation within and among populations. Reproduction is costly in anadromous salmonid fishes. For example, older age at first maturity is linked to greater reproductive success in Atlantic salmon, but comes at the expense of higher mortality risk (e.g. in the ocean) before reproduction. Similarly, most Atlantic salmon individuals adopt a semelparous strategy, in which individuals maximise gonadal investment by also catabolising vital body mass as an energy source, but this most likely comes at the expense mortality at the end of the reproduction event (Jonsson *et al.*, 1997; Fleming & Einum, 2010; Penney & Moffitt, 2013). In contrast, iteroparous individuals as shown in other salmonid species have considerably lower reproductive success in their first reproduction event compared to semelparous individuals, but may attain a higher overall reproductive success, i.e. if the survival in the ocean between two spawning periods is high (see Christie *et al.*, 2018).

Iteroparity and earlier age at first maturity are phenotypically correlated, likely as a result of younger individuals investing less energy in their first reproductive event and hence having better post-reproduction recovery (Jonsson *et al.*, 1991b; 1997; Niemelä *et al.*, 2006a). We therefore hypothesized that these traits are also co-inherited, and we thus employed a gene-trait association approach to test for the existence of pleiotropy between sea age at first maturity and iteroparity in Atlantic salmon. More specifically, we took advantage of a recently reported large-effect genomic

locus controlling age at first maturity in the region of the *vgll3* gene on Atlantic salmon chromosome 25 (Ayllon *et al.*, 2015; Barson *et al.*, 2015) and explored whether the same genome region explained variation in iteroparity.

Materials and Methods

Study site and life history of repeat spawning salmon in the Teno River

Located in far-north Europe (68–70°N, 25–27°E), the Teno River runs between Finland and Norway and drains north into the Barents Sea at the Tana Fjord (Fig. 1). The river supports one of the world's largest wild Atlantic salmon populations and accounts for up to 20% of the riverine Atlantic salmon catches in Europe (ICES 2013). The Teno River supports over 20 Atlantic salmon sub-populations (Vähä *et al.*, 2017) with notably high genetic and life-history variation within and between populations (Vähä *et al.*, 2007; Aykanat *et al.*, 2015; Vähä *et al.*, 2017; Erkinaro *et al.*, 2018). Age at smoltification (i.e., number of years spent in fresh water prior to outward migration to the sea) varies between two and eight years, while the time spent in the marine environment prior to first maturation, called sea age at maturity, varies from one to five years. In addition, a proportion of individuals have an iteroparous life history, in which individuals attempt to reproduce more than once, and sometimes up to three times, in their adult life (Niemelä *et al.*, 2006a; Erkinaro *et al.*, 2018) (see also Figure S1).

In the Teno River, individuals who survive their first spawning event most often spend a full year at sea for conditioning before returning to the river in the following summer (termed as alternate year repeat spawners, Figure S1). Of the repeat-spawning individuals having first spent one, two or three years at sea prior to their first reproductive event, 85% spend a full year at sea for re-conditioning (following the first reproduction event) before returning for a second reproduction attempt (Erkinaro *et al.*, 2018). These three repeat spawning, iteroparous, strategies (i.e. categorised by their

one, two or three year duration at sea prior to first migration) were the focus of our analyses and are referred to as 1S1, 2S1 and 3S1, respectively.

Unlike most Pacific salmon species, Atlantic salmon is categorised as an iteroparous species. Iteroparity is an alternative reproduction strategy in Atlantic salmon with 11% prevalence, on average, ranging from <1% to >40%, (Fleming, 1996). Within the context of this study, individuals were classified as iteroparous if they had spawned previously (see also, Christie *et al.*, 2018). The abundance of a specific iteroparous life history strategy in a cohort is mostly correlated with the abundance of the first-time spawner group in the respective cohort. (Niemelä *et al.*, 2006a; Erkinaro *et al.*, 2018). In the Teno River, catch statistics indicate that the average proportion of repeat-spawner salmon has been 5% over the last four decades, but the abundance and proportion of repeat spawner individuals has been increasing since the early 2000s, which cannot be explained only by the increase in the abundance of the corresponding cohort of first-time spawning fish (Figure S2, Erkinaro *et al.*, 2018).

Sample collection

Atlantic salmon scales can be used to infer a number of life-history characteristics, including smolt age (years spent in fresh water from hatch sea migration), sea age to first spawning event (years spent at sea prior to first spawning, also referred to as age at maturity), and iteroparity (evidence of, and time between, multiple spawning migrations. Erkinaro et al. 2018; Figure S1). The fish samples used in this study were from a scale collection archive of Atlantic salmon from the Teno River system, which spans more than four decades and is maintained by Natural Resources Institute Finland (LUKE; formerly Finnish Game and Fisheries Research Institute, RKTL). Scale samples were composed of riverine catches of anadromous adult Atlantic salmon. These scales have been collected by co-operating, trained fishers of the Teno River system, who have also been recording phenotypic traits (e.g., fish total length, sex) and the location, date and method of capture of the fish. As defined by standard guidelines, scales were sampled from below the adipose fin, just above

147 the lateral line, where the age of the fish can be most reliably inferred (ICES, 2011). The collected
148 scales were later dried and archived in paper envelopes at the Teno River Fisheries Research station
149 of LUKE in Utsjoki, Finland. Trained technicians inferred the freshwater and sea age of fish using
150 standard methods (ICES, 2011).

151 For this study, first-time spawners were defined as adults captured in fresh water whilst returning
152 from the sea for their first reproduction attempt. In total, 643 first-time spawners representing three
153 sea age at first maturity classes were included in this study (N_{1SW}=228, N_{2SW}=170, and N_{3SW}=245,
154 where 1*SW*, 2*SW* and 3*SW* denote the years (sea-winters) spent at sea before the first breeding
155 attempt). The phenotypic and genotypic information of these individuals has previously been
156 reported in (Aykanat *et al.*, 2015). The samples represented individuals captured between 2001 and
157 2003, along a ~130-km stretch of the Teno River mainstem, reaching ca. 210 km from the sea
158 between 2001 and 2003 (see Aykanat *et al.* (2015) for details). Sampled fish were captured in the
159 last four weeks of the fishing season, in August (two to four weeks after most individuals have
160 entered the river), to minimize the number of fish from tributary and headwater populations
161 (Erkinaro *et al.*, 2010). Using the abovementioned dataset, Aykanat *et al.* (2015) identified two sub-
162 populations that have been subsequently identified to represent the Teno mainstem (Tenojoki,
163 referred to as sub-population 1 and Inarijoki (sub-population 2) sub-populations (Pritchard *et al.*,
164 2018).

165 We further studied scales from 492 repeat spawner individuals (N_{1S1}=225, N_{2S1}=155, and N_{3S1}=112)
166 that had good DNA quality. Fish were selected non-randomly with respect to sea age and sex to
167 increase sample size of samples from low frequency life history types. Most fish were captured
168 between 2001 and 2008, but a small proportion of sampled fish were captured during later years
169 (<8% sampled between 2008 and 2014. Table S1). Similar to the first-time spawner sampling
170 scheme, we targeted samples collected later in the season. However, repeat spawner sampling

spanned a broader time window within the fishing season due to the much lower total number of repeat spawner fish late in the season, since repeat spawners tend to return to breeding grounds earlier than the first-time spawners (Niemelä *et al.*, 2006b, Table S1).

DNA extraction, sex determination and SNP genotyping by targeted sequencing

DNA extraction and sex determination of first-time spawners were carried out as described elsewhere (Johnston *et al.*, 2014; Aykanat *et al.*, 2015). For the repeat spawner group, DNA was extracted from one to two scales per individual using a QIAamp 96 DNA QIAcube HT Kit (Qiagen), following the manufacturer's protocol, and with an initial proteinase K digestion step. Quality and concentration of all DNA extractions was assessed using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc.).

Repeat spawner samples were genotyped by targeted sequencing that included 194 SNP loci and the sex determination locus (*sdy*) as outlined in Aykanat *et al.* (2016), with minor modifications. Briefly, genomic regions were first amplified in two multiplex PCR reactions using site-specific primers with adapter sequences. After an SPRI bead clean-up to reduce short, non-specific reads, the PCR products of each individual were combined and re-amplified with adapter-specific primers containing Ion Torrent and sample-specific sequences. The PCR product was again purified by SPRI bead clean-up, quantified with Qubit 2.0 fluorimeter, and pooled in equimolar concentrations into one library (maximum 288 samples together). The pooled library was diluted for template preparation using the Ion PGM Hi-Q OT2 kit for Ion AmpliSeq DNA Library and OT2 for 200 bp reads and enrichment steps (ES) according to the manufacturer’s instructions. Finally, samples were sequenced using an Ion PGM Hi-Q sequencing kit and Ion 318 Chip 2 following the manufacturer’s guidelines.

SNPs in the targeted sequencing panel (N=194) were described in Aykanat *et al.* (2016). The panel consists of putatively neutral and highly diverged SNPs between the Inarijoki and Tenojoki sub-populations (neutral module: N=136, outlier module: N=53) and potentially functionally important SNPs that are associated with sea age at maturity on chromosomes 9 and 25 (sea age module, N=5). The baseline and outlier modules allow for the quantification of population genetic parameters and reliably assign population of origin (Aykanat *et al.* (2016). The sea age module consists of four SNPs on chromosome 25, located in the genome region associated with age at maturity (Ayllon *et al.*, 2015; Barson *et al.*, 2015), of which two are missense SNPs in *vgll3* ($vgll3_{Met54Thr}$ and $vgll3_{Asn323Lys}$), one is a missense SNP in *akap11* ($akap11_{Val214Met}$), and one is the SNP with the most significant association with age at maturity (i.e., $vgll3_{TOP}$ from Barson *et al.*, 2015). In addition, the module has an SNP from the chromosome 9 region that exhibits a strong association with sea age at maturity prior to population structure correction (Barson *et al.*, 2015). This SNP is 34.5 kb away from and in complete linkage disequilibrium with the $SIX6_{TOP}$ SNP from Barson *et al.* (2015) (termed $SIX6_{TOP.LD}$ here). The majority of the SNP data for first-time spawners were taken from previous studies (Johnston *et al.*, 2014; Aykanat *et al.*, 2015; Barson *et al.*, 2015), whereby the genotypoing was perfomred using 5568 SNP loci using a custom-designed Illumina® iSelect SNP array (Johnston *et al.*, 2014; Aykanat *et al.*, 2015) or a custom 220,000 SNP Affymetrix Axiom array (Barson *et al.*, 2015). Two missense SNPs, $vgll3_{Met54Thr}$ and $vgll3_{Asn323Lys}$, which were not scored in those arrays were sequenced using the Ion Torrent platform as described above.

Unless otherwise noted, all statistical analyses were performed using R software v.3.2.5 (R Core Team 2013). Raw *fastq* output from the Ion Torrent server was scored using custom R scripts as outlined in Aykanat *et al.* (2016). Briefly, forward and reverse barcodes were used as identifiers to assign reads to individuals, followed by assigning within-individual data to each locus by matching reads to locus-specific primers (allowing for one mismatch). Reads were further refined by a sequence pattern match, this time to a 9-bp region surrounding each SNP locus, without a mismatch

allowed. Finally, loci with coverage less than 13 were excluded, and the remainder were assigned a genotype as in Campbell *et al.* (2015). After excluding three SNPs with low genotyping rates (<75%), genotyping success on average was 96.5%. Molecular sexing was conducted by estimating read counts of the *sdy* gene (normalized to mean read count for every individual) using an arbitrary threshold as outlined in Aykanat *et al.* (2016).

Genetic assignment of individuals using genetic baseline data

All sampled fish were captured as late in the season as possible in order to maximize the proportion of fish from focal populations (i.e., Tenojoki and Inarijoki). Due to the lower number of iteroparous individuals in the population, repeat-spawner individuals captured earlier in the season were included, which may have increased the proportion of fish from non-focal, tributary populations amongst this group (Table S1). Therefore, all samples underwent population assignment to enable exclusion of individuals potentially originating from non-focal populations using 175 non-sea-age associated SNPs that were successfully genotyped both in the dataset and in baseline samples. The baseline consisted individuals from 23 sub-populations within the Teno system (Fig. 1). We first estimated the allele frequency distributions of the SNPs across the baseline populations and then estimated the likelihood of each individual to originate from each of those populations using a frequency-based method (see Table S2 for more details). The robustness of assignments was shown to be high using simulations, whereby random generation of 1000 genotypes per population showed high true assignment rates to Tenojoki and Inarijoki (93.2% and 91.2%, respectively); a negligible proportion of individuals from other populations were incorrectly assigned to these focal populations with high confidence (0.3% and 0.8%, respectively).

Out of 643 first-time spawners, 335 (52.1%) and 167 (26.0%) were assigned to Tenojoki and Inarijoki, respectively (78.1% in total), while 16.6% were unassigned, and 5.3% were assigned to other populations (Table S2). For 478 repeat-spawning individuals, 115 and 141 fish were assigned

to Tenojoki and Inarijoki populations, respectively (53.6%, in total), leaving 137 fish (28.7%) unassigned and 85 fish (17.8%) confidently assigned to other populations in the Teno system (Table S2). The high proportion of misassigned fish in the repeat spawning category was expected, given that this group also included individuals sampled earlier in the fishing season (May-July). Individuals from these months included a higher proportion of fish captured in their non-natal spawning grounds. In contrast, the proportion of fish assigned to non-focal populations was similar between first-time spawners and repeat spawners in August (exact-test, $p>0.05$, Table S3**)**. Following this assignment procedure, the final data set included 502 first-time spawners and 256 repeat spawners (Table S2).

Testing genetic associations to iteroparity

In Atlantic salmon, sea age at first maturation (i.e., sea age) is strongly controlled by one genomic region (i.e., around the *vgll3* gene, approximately 28.65 Mb on chromosome 25, Barson *et al.*, 2015). An additional region, around the *six6* gene on chromosome 9, also exhibits a strong association with sea age at the population level, though the signal diminishes after correcting for structure (Barson *et al.*, 2015). Here, we tested if variation at five SNPs linked to sea age at first maturity on chromosome 9 and 25 ($vgll3_{TOP}$, $vgll3_{Met54Thr}$, $vgll3_{Asn323Lys}$, $akap11_{Val214Met,}$ $six6_{TOP}$ (see above) were also linked to the repeat spawning life history. To do this, we first identified the most parsimonious null model that fit the data without genetic effects and then employed the genetic model over it. This allowed for us to avoid any biased inference that may have emerged as a result of unequal proportions of previous spawners within sea age groups, sex, populations or their higher-order interactions. We employed a generalised linear model with a binomial error structure, where repeat-spawner fish were coded as a Boolean variable “1”, first-time spawners were coded as “0”, and the allelic effect was an independent variable. Origin of population, sea age at first maturity, and sex were included in the model as cofactors. Using a semi-automatic model-scanning approach

in the MuMIn package (Barton, 2018) in R (R Core Team 2013), a full-null model (all cofactors with all possible interactions, but without the genetic term) and reduced null models were compared by the corrected Akaike information criterion score (*AICc*), which is an *AIC* score with a stronger penalty for complex models. The null model explaining the data the best (with the lowest *AICc* score) was then used as the null hypothesis when testing the genetic effect. The genetic effect was then included in the optimum null model as an additional independent variable. As above, alternative genetic models with all possible interactions of the genetic term with variables in the the null model were tested, in which the genetic term were included in the model either as numerical (i.e., to model additive genetic inheritance) or as categorical effect (i.e., to test if any non-additive models explained the data better than the additive model). Based on the prevalence of iteroparity in the populations, a small proportion of first-time spawners in the dataset are likely to have spawned a second time hadn't they captured by fishers (~5% of the samples), making the statistical analyses between two groups conservative.

Repeat spawner sampling spanned a broader time window than the first-time spawner sampling due to the much smaller total number of repeat-spawner fish late in the season (Niemelä *et al.*, 2006b, Table S1). We therefore explored if a mismatch in sampling timing between first-time and repeat spawners could have resulted in a spurious genetic association between traits, by employing a linear model in which the date of sampling (i.e. date of capture as number of days in a Julian calendar) on genotype, nested within each sea-age group. The sensitivity of the genetic association to the timing of migration was further assessed by permutation by resampling with replacement (with sample sizes equal to actual sample size) within different periods in the sampling season. The periods of resampling included i) only August, to echo the first-time spawner sampling period, ii) the mean and standard deviation of the natural distribution of timing of repeat-spawner capture, and iii) resampling from all individuals in the dataset for direct comparison of distributions between restricted resampling and the full one.

Finally, to understand possible growth variation among genotypes associated with repeat spawning, length at capture was investigated as a function of genotype within each life history group (i.e., repeat and first-time spawners). We again employed a model evaluation approach as above, where a full-interaction model, consisting of sea age at first maturity, sex, and population, and reduced models were evaluated to find the most parsimonious model. The response variable, total length at capture, was log-transformed.

Results

The best null model explaining the repeat-spawning life-history strategy (without genetic effects as a term) included population, sea age at first spawning, sex and the interaction between sex and population as factors (Table S4). The highest improvement in the fit of the model for explaining repeat spawning compared to the null model was obtained when SNPs in the chromosome 25 region associated with sea age at maturity were included in the model (Fig. 2). When alternative parameterizations of genetic effects were evaluated for goodness of fit across those SNPs in the region (i.e. by only including individuals that were scored for all four SNPs), the missense polymorphism at $vgll3_{Met54Thr}$ exhibited the highest fit across all loci (Table S5). In the best-supported model, the genetic architecture was consistent with non-additivity (SNP was coded as a categorical effect), which was dependent on sex as well as sea age at maturity (Table 1, Table S5, Table S6). Even though $vgll3_{Met54Thr}$ explained the data better than other loci in the chromosome 25 region, the difference was not substantial. For example, when the dataset was resampled 1000 times (random sampling within sex and sea age groups with replacement), models including $vgll3_{Met54Thr}$ were the best-fitting models in 42% of cases, while models with $vgll3_{TOP}$, $vgll3_{Asn323Lys}$, and $akap11_{Val214Met}$ were better fitting 21.9%, 14.9% and 21.2% of the time, respectively. Therefore, we cannot clearly distinguish between the importance of the SNPs in the region with respect to their association with repeat spawner prevalence.

315 The odds ratio of survival to second spawning was estimated as a function of *vgll3*$_{Met54Thr}$ genetic
316 variation using the most parsimonious model. Generally, the genotype associated with earlier first
317 maturity at sea (*EE*, *E* denotes the allele associated with early first-time maturation) was more likely
318 to survive to second spawning than the genotype associated with later first-time maturation (*LL*, *L*
319 denotes the allele associated with late first-time maturation) (Fig. 3, Table 1). This effect was age-
320 dependent with a slight but not significant effect of population and a sex × population interaction
321 (Fig. 3, Table 1, Table S7). The odds ratio of survival to second spawning differed less between the
322 genotypes in 1S1 fish: the *EE* genotype was 1.19 (0.91-1.91, 90% CI, p=0.14) and 1.24 (1.00-1.57,
323 90% CI, p=0.051) times more likely to survive to second spawning compared to the *LL* and *EL*
324 genotypes, respectively. Older sea age groups exhibited stronger differences between genotypes,
325 where the odds ratio between *LL* and *EE* genotype fish was 1.50 (1.12-2.05, 90% CI, p=0.012) and
326 2.03 (1.07-3.57, 90% CI, p=0.035) for 2S1 and 3S1 fish, respectively (Fig. 3, Table S7). The
327 genetic architecture was consistent with dominance, which, however, was age-dependent (Fig. 3,
328 Table S7). In 3S1 fish, the genetic architecture significantly deviated from additivity (p=0.014,
329 Table S7), whereby the *L* allele was dominant, while in 1S1 fish, the genetic architecture marginally
330 but not statistically deviated from additivity in the reverse direction. In agreement with this, a model
331 in which *vgll3*$_{Met54Thr}$ genotype modelled with complete dominant genetic architecture and age-
332 dependently (i.e. as reflected in Fig. 3) had the best model fit among other alternative
333 parameterisation of dominance (Table S8).

334 We tested the potential bias that the mismatch in the date of sampling between first-time and repeat
335 spawners may exert to the estimation of the genetic association. The *vgll3*$_{Met54Thr}$ locus was a
336 marginal predictor of date of sampling in a season (p=0.054, *E* allele was associated with 4.1 days
337 ±2.2 SE later sampling in a season, see: Table S9). Permutation by re-sampling along different
338 sampling periods (10000 times) suggested that the genetic association would have been much more
339 pronounced if all repeat spawner individuals were sampled in the same period as the first-time

spawners, in August (Figure S3). Conversely, if sampling had reflected the natural distribution of repeat spawners, we predicted only a slight, non-significant decrease in the model fit (Figure S3), suggesting that our results are relatively robust to sampling time. Day of sampling was not a significant predictor of genotype in the first-time spawner group (data not shown).

Finally, we tested for the allelic effect of $vgll3_{Met54Thr}$ on total length at capture separately for first-time and repeat spawner groups. In the first-time spawners, the most parsimonious model included the SNP effect nwithout any higher-order interactions, and the *L* allele was positively associated with length (Table 2, Table S10). This relationship was similar to that reported by Barson *et al.* (2015) for the $vgll3_{TOP}$ locus. In contrast, there was no allelic association with length in the repeat spawner group (Table 2, Table S10). These two results combined suggest that allele-specific growth differences were offset in the repeat-spawner fish.

Discussion

Survival following the first breeding event to potentially reproduce additional times (i.e., repeat spawning) is an important life history feature of Atlantic salmon. The trait is closely connected to fitness. Hence, understanding the ecological dynamics around the trait variation and its genetic underpinnings is important for predicting Atlantic salmon survival in the wild. Here, we showed that survival until the second spawning event was linked to the genome region that is also associated with sea age at maturity (Ayllon *et al.*, 2015; Barson *et al.*, 2015), another important fitness-linked life history trait. The allele associated with a higher prevalence of repeat spawning was also linked to earlier first-time maturation (i.e., the *E* allele in Barson *et al.*, 2015). This genetic correlation suggests that individuals with the *E* allele increase their fitness (relative to the *L* allele, associated with late maturation) not only by means of spending less time at sea, hence lowering the risk of mortality and shorter generation time, but also by increasing reproductive success by participating in multiple spawning events.

364 The genetic linkage between an iteroparous reproductive strategy and earlier first-time spawning is
365 in accordance with energy allocation similarities between these two life history strategies as
366 predicted by life history theory: investment in reproduction is higher in semelparous
367 organisms/individuals (a single reproduction event before death) compared to those with and an
368 iteroparous reproductive strategy (repeat spawning in this case, Crespi & Teo, 2002). Reproductive
369 investment is indeed higher in large, late-maturing individuals who have spent two or more years at
370 sea prior to first spawning, as quantified by a higher proportion of energy deposited to gonadal
371 organs (Jonsson *et al.*, 1997; Jonsson & Jonsson, 2003), while gonadal investment in repeat
372 spawners is lower (Fleming & Einum, 2010). Furthermore, the *vgll3* gene, the most likely gene
373 candidate in the region associated with both traits, regulates adipose content and body weight in
374 mice (Halperin *et al.*, 2013), both of which are postulated to regulate maturation age and repeat
375 spawning in Atlantic salmon (Friedland & Haas, 1996; Jonsson *et al.*, 1997; Taranger *et al.*, 2010).
376 Our results support the notion that the genetic variation in these two life history strategies may be
377 mediated by similar a physiological cascade controlled, at least in part, by the same genomic region.

378 Even though sea age at maturity and repeat spawning are genetically linked and regulated by similar
379 physiological cascades, it is unclear if the same SNP is linked to trait variation in both traits (i.e.,
380 pleiotropy in the strict sense, see Wagner and Zhang, 2011) or if the association is a result of a
381 causal SNPs being tightly linked in the region. Although $vgll3_{Met54Thr}$, a SNP causing a potentially
382 functionally important missense substitution in the *vgll3* gene (Barson *et al.*, 2015), appears to be a
383 prime candidate for a causal association to iteroparity, other candidate SNPs in the region were not
384 conclusively ruled out as the region of highest association. Likewise, causality between SNPs in the
385 region and age at maturity is yet to be tested further (Barson *et al.*, 2015). Narrowing down the
386 causal SNP will be important to gain a better understanding of the physiological mechanisms
387 underlying the trait variation, and in this particular case, to pinpoint the nature of the linkage
388 between the two traits. This may be particularly important in estimating potential evolutionary

389 trajectories of life history co-evolution in salmon (Etterson & Shaw, 2001; Steppan *et al.*, 2002;
390 Conner *et al.*, 2011). For example, true pleiotropy may be slower to break up than genetic
391 correlations caused by physical linkage, e.g., if the genetic linkage is maladaptive (Mackay, 2001;
392 Gardner & Latta, 2007; Mackay *et al.*, 2009; Paaby & Rockman, 2013). The nature of the linkage
393 should be further investigated to better evaluate the co-evolutionary potential between life history
394 traits over contemporary time scales (Gardner & Latta, 2007). On the other hand, distinguishing
395 true pleiotropy from tight genetic linkage is demanding, and may require substantial effort to
396 elucidate the whole extent of polymorphisms in the genome region, and to validate the causality,
397 e.g. perhaps by means of genetic editing methodologies. Such research is still impractical in non-
398 model organisms.

399 Our framework does not quantify the extent of genetic correlation between the two traits, e.g. by the
400 use of multivariate modelling framework. However, the prevalence of genetic linkage in both traits
401 across a narrow genomic region intrinsically implies its presence, thus making multi-trait
402 evolutionary inferences feasible. In most documented cases, the evolutionary potential of
403 genetically correlated fitness traits is constrained by the opposing directional response of correlated
404 trait values to the selection gradient (e.g. Etterson & Shaw, 2001; Charmantier *et al.*, 2006;
405 Theriault *et al.*, 2007; Fordyce & Nice, 2008; Nussey *et al.*, 2008; Simon *et al.*, 2016), which is
406 predicted to help maintain variation in fitness traits (Roff, 1996). In our study, the trait variation
407 within a population was not maintained by opposing selective directions of correlated trait values,
408 since sea age at maturity is not under directional selection, but variation is being maintained by
409 balancing selection between younger and older maturation ages (Barson *et al.*, 2015). Therefore, the
410 dynamics underlying the genetic covariation cannot be explained by antagonistic trait correlation.
411 Assuming that iteroparity in Atlantic salmon is linked to higher reproductive success, the genetic
412 basis of trait co-variation between an iteroparous reproductive strategy and sea age at maturity
413 would reinforce the prevalence of a younger age structure; hence, a younger population age

414 structure would be predicted as opposed to a case where the genetic correlation were not accounted
415 for. On the other hand, it is not clear if repeat spawning is always linked to higher fitness. In
416 steelhead salmon (*Oncorhynchus mykiss*), iteroparity is linked to increased overall reproductive
417 success, but the first-time reproductive output of repeat spawners is less than that of semelparous
418 individuals at the same age (Christie *et al.*, 2018), suggesting that a fitness advantage may be
419 reversed if conditions promoting post-spawning survival at sea deteriorate (e.g., Chaput & Benoit,
420 2012).

421 There was a contradiction between the sex-specific phenotypic patterns of repeat spawning
422 compared to the observed genetic basis of repeat spawning variation (Table 1). Females tend to
423 have a higher level of iteroparity than males (Niemelä *et al.*, 2006a), but this phenotypic difference
424 was not translated into a genetic association at the $vgll3_{Met54Thr}$ locus, which is independent of sex.
425 This lends support to the notion that the sex-specific differences in repeat spawning may be
426 attributed to sex-specific behavioural differences rather than physiological differences (e.g.,
427 Niemelä *et al.*, 2006a). For example, despite the much lower gonadosomatic index of male gonads
428 (Jonsson *et al.*, 1991b; 1997), males are more aggressive and more likely to attempt reproduction
429 until exhausting themselves to the point of death and/or sustain more damage during reproduction
430 (Jonsson *et al.*, 1997), while female reproductive effort is limited to egg deposition. From this
431 perspective, our results are consistent with a notion that genetic variation in the $vgll3_{Met54Thr}$ locus is
432 associated with physiological mechanisms that are not linked to sex-specific differences in
433 behaviour. On the other hand, the genetic regulation in repeat spawning prevalence was
434 substantially stronger in later-maturing fish. Stronger physiological constraints dramatically impede
435 survival in older sea age groups, likely as a result of higher energetic loss during maturation and
436 spawning, as well as relatively higher maintenance metabolism (Jonsson *et al.*, 1991a; Jonsson *et*
437 *al.*, 1997; Metcalfe *et al.*, 2015). The co-inheritance between maturation age and repeat spawning
438 prevalence suggests the same energetic constraints underlying the physiological basis of both traits

439 may be controlled by the *vgll3* locus. Overall, understanding the interplay between the phenotypic
440 and genetic correlation between repeat spawning and other phenotypic indices (sex and sea age at
441 maturity) will provide further insights into the functional basis of the association between these
442 traits.

443 Post-smolt growth at sea is an important determinant of survival in salmon (e.g., Friedland *et al.*,
444 2005) and has been shown to be positively correlated with early age at maturity (Friedland & Haas,
445 1996). Likewise, better growth conditions at sea are correlated to post-spawning survival (Chaput &
446 Benoit, 2012). Here, unlike the first-time spawners, we showed that size did not differ between
447 repeat spawners with different *vgll3* genotypes (Table 2). However, it is unclear from the terminal
448 length data at which life history stage the genotype-specific size differences observed in first-time
449 spawners were offset in repeat spawners. Two plausible explanations exist. i) Survival after the first
450 spawning event is length-dependent for individuals with the *EE* genotype but not for those with the
451 *LL* genotype. ii) Fish with the *EE* genotype grow faster than *LL* individuals after the first spawning
452 event. By measuring the length of surviving post-spawned fish in the Teno River prior to their
453 second sea migration, Niemelä *et al.* (2000) showed that repeat spawners were, on average, larger
454 than first-time spawners. This information is consistent with the first possibility, that larger fish
455 with the *EE* genotype are likely to survive the first spawning, while this survival is length
456 independent for fish with the *LL* genotype.

457 To investigate the second possibility further, i.e. if post-spawning growth is greater in *EE*
458 individuals than in *LL* individuals, we performed a *post hoc* analysis and quantified post-spawning
459 scale growth as a proxy for post-spawning growth, using a similar model selection framework as for
460 other analyses (Appendix S1, Figure S4). This analysis appeared to be underpowered, as many
461 model structures, including null models, exhibited similarly high explanatory powers ($\Delta AICc<2$ in
462 the most parsimonious models, see Table S11), As such, we could not refute the null hypothesis and

conclude that there is insufficient evidence to support better post-spawn growth associated with the *EE* genotype. Interestingly, however, in all non-null plausible models the estimated SNP effect was in a contrasting direction to the expectation (*AICc* weight = 73.4%), lending no support to a scenario where better post-spawn growth in the *EE* genotype may explain growth patterns in repeat spawners. Finally, in all plausible models, post-spawning growth was negatively correlated to sea age (Table S11), suggesting that older sea age groups had limited post-spawning growth. In fact, when post- and pre-spawning scale growth were plotted over total length separately, a correlation between scale growth after first spawning and total length was only observed in 1S1 fish. This further illustrates that the post-spawn growth is dependent on sea age (Figure S5, see also Supplementary Information). Overall, this *ad hoc* analysis confirmed that the genotype-dependent differences observed between repeat spawners and first-time spawners arose prior to first migration, rather than after. In addition, we further showed growth during the post-spawning period was limited in later-maturing fish. This was probably as a result of the higher energy demands of larger, older maturing fish to restore and maintain their body mass after a reproduction event (Jonsson *et al.*, 1997). Individuals maturing later invest more energy in reproduction (Jonsson *et al.*, 1997), and therefore, it might be more difficult to maintain positive energy balance after the first reproduction event. Relatively little is known on the dynamics of iteroparity in Atlantic salmon (Thorstad *et al.*, 2010), and our results thus provide valuable insight on the growth dynamics of alternative reproductive tactics.

The abundance of Atlantic salmon populations has declined substantially over the past 40 years and is at an all-time low level (ICES; Chaput, 2012). This decline is coupled with demographic changes, mostly towards younger age structure (Chaput, 2012; Czorlich *et al.*, 2018), and global climate change will likely affect population structure further (Friedland *et al.*, 2005; Friedland *et al.*, 2009; Hedger *et al.*, 2013; Mills *et al.*, 2013; Piou & Prevost, 2013; Jonsson *et al.*, 2016). Many salmon management regimes consider the repeat spawner phenotype a suitable substitute for dwindling

numbers of large, late-maturing, first-time spawning individuals, and their presence in populations has been linked with improved genetic stability (Hatch *et al.*, 2004; Niemelä *et al.*, 2006a; Narum *et al.*, 2008; Seamons & Quinn, 2009; Chaput & Benoit, 2012; Reid & Chaput, 2012). However, the genetic evidence presented here suggests that increases in repeat spawner numbers may be associated with decreasing age structure of first-time spawners, and directional selection changing sea age composition (e.g., in response to environmental changes) may result in changes in repeat spawner composition, which may influence the demographic structure further. Despite the fact that genetic effects have been increasingly included in demographic models (i.e., demo-genetic models, i.e., (Dunlop *et al.*, 2009; Reed *et al.*, 2011; Piou & Prévost, 2012), accurately accounting for genetic architecture, e.g., (Kuparinen & Hutchings, 2017), gene-trait association, genetic correlations, and functional links between correlated traits would provide more realistic predictions, which could improve future conservation and management efforts.

References

Agrawal, A.F. & Stinchcombe, J.R. 2009. How much do genetic covariances alter the rate of adaptation? *Proc. Biol. Sci.* **276:** 1183-1191.

Albert, A. Y., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., Balabhadra, S., Kingsley, D. M. & Schluter, D. 2008. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* **62:** 76-85.

Arnold, S. J. 1981. Behavioral variation in natural populations. I. Phenotypic, genetic and environmental correlations between chemoreceptive responses to prey in the garter snake, *Thamnophis elegans*. *Evolution* **35**: 489-509.

Aykanat, T., Johnston, S.E., Orell, P., Niemela, E., Erkinaro, J. & Primmer, C.R. 2015. Low but significant genetic differentiation underlies biologically meaningful phenotypic divergence in a large Atlantic salmon population. *Mol. Ecol.* **24:** 5158-5174.

Aykanat, T., Lindqvist, M., Pritchard, V.L. & Primmer, C.R. 2016. From population genomics to conservation and management: a workflow for targeted analysis of markers identified using genome-wide approaches in Atlantic salmon *Salmo salar*. *J. Fish. Biol.* **89**: 2658-2679.

Ayllon, F., Kjaerner-Semb, E., Furmanek, T., Wennevik, V., Solberg, M.F., Dahle, G., Taranger, G.L., Glover, K.A., Almen, M.S., Rubin, C.J., Edvardsen, R.B. & Wargelius, A. 2015. The vgll3 locus controls age at maturity in wild and domesticated Atlantic salmon (*Salmo salar* L.) males. *PLoS Genet.* **11:** e1005628.

Barson, N.J., Aykanat, T., Hindar, K., Baranski, M., Bolstad, G.H., Fiske, P., Jacq, C., Jensen, A.J., Johnston, S.E., Karlsson, S., Kent, M., Moen, T., Niemela, E., Nome, T., Naesje, T.F., Orell, P., Romakkaniemi, A., Saegrov, H., Urdal, K., Erkinaro, J., Lien, S. & Primmer, C.R. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* **528:** 405-408.

Berenos, C., Ellis, P.A., Pilkington, J.G. & Pemberton, J.M. 2014. Estimating quantitative genetic parameters in wild populations: a comparison of pedigree and genomic approaches. *Mol. Ecol.* **23:** 3434-3451.

Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., ReproGen, C., Psychiatric Genomics, C., Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control, C., Duncan, L., Perry, J.R., Patterson, N., Robinson, E.B., Daly, M.J., Price, A.L. & Neale, B.M. 2015. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47:** 1236-1241.

Campbell, N.R., Harmon, S.A. & Narum, S.R. 2015. Genotyping-in-Thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Mol. Ecol. Resour.* **15:** 855-867.

Careau, V., Reale, D., Humphries, M.M. & Thomas, D.W. 2010. The pace of life under artificial selection: personality, energy expenditure, and longevity are correlated in domestic dogs. *Am. Nat.* **175:** 753-758.

Carlson, S.M. & Seamons, T.R. 2008. A review of quantitative genetic components of fitness in salmonids: implications for adaptation to future change. *Evol. Appl.* **1:** 222-238.

Chaput, G. 2012. Overview of the status of Atlantic salmon (*Salmo salar*) in the North Atlantic and trends in marine mortality. *Ices J. Mar. Sci.* **69:** 1538-1548.

Chaput, G. & Benoit, H.P. 2012. Evidence for bottom-up trophic effects on return rates to a second spawning for Atlantic salmon (*Salmo salar*) from the Miramichi River, Canada. *Ices J. Mar. Sci.* **69:** 1656-1667.

Charmantier, A., Perrins, C., McCleery, R.H. & Sheldon, B.C. 2006. Quantitative genetics of age at reproduction in wild swans: Support for antagonistic pleiotropy models of senescence. *Proc. Natl. Acad. Sci. USA* **103:** 6587-6592.

548 Chen, C.Y., Misztal, I., Aguilar, I., Tsuruta, S., Meuwissen, T.H., Aggrey, S.E., Wing, T. & Muir,
549 W.M. 2011. Genome-wide marker-assisted selection combining all pedigree phenotypic
550 information with genotypic data in one step: An example using broiler chickens. *J. Anim.*
551 *Sci.* **89:** 23-28.
552 Chevin, L.M., Martin, G. & Lenormand, T. 2010. Fisher's model and the genomics of adaptation:
553 restricted pleiotropy, heterogenous mutation, and parallel evolution. *Evolution* **64:** 3213-
554 3231.
555 Chirgwin, E., Monro, K., Sgro, C.M. & Marshall, D.J. 2015. Revealing hidden evolutionary
556 capacity to cope with global change. *Glob. Chang. Biol.* **21:** 3356-3366.
557 Christie, M.R., McNickle, G.G., French, R.A. & Blouin, M.S. 2018. Life history variation is
558 maintained by fitness trade-offs and negative frequency-dependent selection. *Proc Natl*
559 *Acad Sci U S A*.
560 Clements, M.N., Clutton-Brock, T.H., Guinness, F.E., Pemberton, J.M. & Kruuk, L.E. 2011.
561 Variances and covariances of phenological traits in a wild mammal population. *Evolution*
562 **65:** 788-801.
563 Conner, J.K., Karoly, K., Stewart, C., Koelling, V.A., Sahli, H.F. & Shaw, F.H. 2011. Rapid
564 independent trait evolution despite a strong pleiotropic genetic correlation. *Am Nat* **178:**
565 429-441.
566 Crespi, B.J. & Teo, R. 2002. Comparative phylogenetic analysis of the evolution of semelparity and
567 life history in salmonid fishes. *Evolution* **56:** 1008-1020.
568 Czorlich, Y., Aykanat, T., Erkinaro, J., Orell, P. & Primmer, C. R. 2018. Rapid sex-specific
569 evolution of age at maturity is shaped by genetic architecture in Atlantic salmon. *Nat. Ecol.*
570 *Evol.* **2**: 1800-1807.
571 DeLong, J. P. 2017. Ecological pleiotropy suppresses the dynamic feedback generated by a rapidly
572 changing trait. *Am. Nat.* **189**: 592-597.
573 Dunlop, E.S., Heino, M. & Dieckmann, U. 2009. Eco-genetic modeling of contemporary life-
574 history evolution. *Ecol. Appl.* **19:** 1815-1834.
575 Duputie, A., Massol, F., Chuine, I., Kirkpatrick, M. & Ronce, O. 2012. How do genetic correlations
576 affect species range shifts in a changing environment? *Ecol. Lett.* **15:** 251-259.
577 Erkinaro, J., Niemelä, E., Vähä, J.-P., Primmer, C.R., Brørs, S. & Hassinen, E. 2010. Distribution
578 and biological characteristics of escaped farmed salmon in a major subarctic wild salmon
579 river: implications for monitoring. *Can. J. Fish. Aquat. Sci.* **67:** 130-142.
580 Erkinaro, J., Czorlich, Y., Orell, P., Kuusela, J., Falkegård, M., Länsman, M., Pulkkinen, H.,
581 Primmer, C.R. & Niemelä, E. 2018. Life history variation across four decades in a diverse
582 population complex of Atlantic salmon in a large subarctic river. *Can. J. Fish. Aquat. Sci.*.
583 Etterson, J.R. & Shaw, R.G. 2001. Constraint to adaptive evolution in response to global warming.
584 *Science* **294:** 151-154.
585 Fleming, I.A. 1996. Reproductive strategies of Atlantic salmon: ecology and evolution. *Rev. Fish.*
586 *Biol. Fisher.* **6:** 379-416.
587 Fleming, I.A. & Einum, S. 2010. Reproductive Ecology: A Tale of Two Sexes. In: *Atlantic Salmon*
588 *Ecology*, pp. 33-65. Wiley-Blackwell.
589 Fordyce, J.A. & Nice, C.C. 2008. Antagonistic, stage-specific selection on defensive chemical
590 sequestration in a toxic butterfly. *Evolution* **62:** 1610-1617.
591 Friedland, K.D. & Haas, R.E. 1996. Marine post-smolt growth and age at maturity of Atlantic
592 salmon. *J. Fish. Biol.* **48:** 1-15.
593 Friedland, K.D., Chaput, G. & MacLean, J.C. 2005. The emerging role of climate in post-smolt
594 growth of Atlantic salmon. *Ices J. Mar. Sci.* **62:** 1338-1349.
595 Friedland, K.D., MacLean, J.C., Hansen, L.P., Peyronnet, A.J., Karlsson, L., Reddin, D.G.,
596 Maoileidigh, N.O. & McCarthy, J.L. 2009. The recruitment of Atlantic salmon in Europe.
597 *Ices J. Mar. Sci.* **66:** 289-304.

598 Gardner, K.M. & Latta, R.G. 2007. Shared quantitative trait loci underlying the genetic correlation
599 between continuous traits. *Mol. Ecol.* **16:** 4195-4209.
600 Gratten, J. & Visscher, P.M. 2016. Genetic pleiotropy in complex traits and diseases: implications
601 for genomic medicine. *Genome Med.* **8:** 78.
602 Griswold, C.K. & Whitlock, M.C. 2003. The genetics of adaptation: the roles of pleiotropy,
603 stabilizing selection and drift in shaping the distribution of bidirectional fixed mutational
604 effects. *Genetics* **165:** 2181-2192.
605 Halperin, D.S., Pan, C., Lusis, A.J. & Tontonoz, P. 2013. Vestigial-like 3 is an inhibitor of
606 adipocyte differentiation. *J. Lipid. Res.* **54:** 473-481.
607 Hammerschlag, A.R., Stringer, S., de Leeuw, C.A., Sniekers, S., Taskesen, E., Watanabe, K.,
608 Blanken, T.F., Dekker, K., Te Lindert, B.H.W., Wassing, R., Jonsdottir, I., Thorleifsson, G.,
609 Stefansson, H., Gislason, T., Berger, K., Schormair, B., Wellmann, J., Winkelmann, J.,
610 Stefansson, K., Oexle, K., Van Someren, E.J.W. & Posthuma, D. 2017. Genome-wide
611 association analysis of insomnia complaints identifies risk genes and genetic overlap with
612 psychiatric and metabolic traits. *Nat. Genet.* **49:** 1584-1592.
613 Hatch, D.R., Branstetter, R.D., Whiteaker, J., Blodgett, J., Bosch, B., Fast, D. & Newsome, T. 2004.
614 Kelt Reconditioning: A Research Project to Enhance Iteroparity in Columbia Basin
615 Steelhead (*Oncorhynchus mykiss*). Annual Report to U.S. Department of Energy, Bonneville
616 Power Administration, Project No. 2000-017. Portland, OR: Bonneville Power
617 Administration.
618 Hedger, R.D., Naesje, T.F., Fiske, P., Ugedal, O., Finstad, A.G. & Thorstad, E.B. 2013. Ice-
619 dependent winter survival of juvenile Atlantic salmon. *Ecol. Evol.* **3:** 523-535.
620 Hellmann, J.J. & Pineda-Krch, M. 2007. Constraints and reinforcement on adaptation under climate
621 change: Selection of genetically correlated traits. *Biological Conservation* **137:** 599-609.
622 Houle, D. 1991. Genetic covariance of fitness correlates - What genetic correlations are made of
623 and why it matters. *Evolution* **45:** 630-648.
624 ICES. 2011. Report of the Workshop on Age Determination of Salmon (WKADS), 18-20 January
625 2011, Galway, Ireland. *ICES CM 2011/ACOM*:44. 67 pp.
626 ICES. 2018. Report of the Working Group on North Atlantic Salmon (WGNAS), 4–13 April 2018,
627 Woods Hole, MA, USA. *ICES CM 2018/ACOM*:21. 386 pp.
628 Johnston, S.E., Orell, P., Pritchard, V.L., Kent, M.P., Lien, S., Niemela, E., Erkinaro, J. & Primmer,
629 C.R. 2014. Genome-wide SNP analysis reveals a genetic basis for sea-age variation in a
630 wild population of Atlantic salmon (*Salmo salar*). *Mol. Ecol.* **23:** 3452-3468.
631 Jonsson, B., Jonsson, N. & Albretsen, J. 2016. Environmental change influences the life history of
632 salmon *Salmo salar* in the North Atlantic Ocean. *J Fish Biol.*
633 Jonsson, N., Hansen, L.P. & Jonsson, B. 1991a. Variation in age, size and repeat spawning of adult
634 atlantic salmon in relation to river discharge. *J. Anim. Ecol.* **60:** 937-947.
635 Jonsson, N., Jonsson, B. & Hansen, L.P. 1991b. Energetic cost of spawning in male and female
636 Atlantic salmon (*Salmo salar* L). *J. Fish. Biol.* **39:** 739-744.
637 Jonsson, N., Jonsson, B. & Hansen, L.P. 1997. Changes in proximate composition and estimates of
638 energetic costs during upstream migration and spawning in Atlantic salmon *Salmo salar*. *J.*
639 *Anim. Ecol.* **66:** 425-436.
640 Jonsson, N. & Jonsson, B. 2003. Energy allocation among developmental stages, age groups, and
641 types of Atlantic salmon (*Salmo salar*) spawners. *Can. J. Fish. Aquat. Sci.* **60:** 506-516.
642 Kadarmideen, H.N., Thompson, R., Coffey, M.P. & Kossaibati, M.A. 2003. Genetic parameters and
643 evaluations from single- and multiple-trait analysis of dairy cow fertility and milk
644 production. *Livestock Production Science* **81:** 183-195.
645 Kause, A., Quinton, C., Airaksinen, S., Ruohonen, K. & Koskela, J. 2011. Quality and production
646 trait genetics of farmed European whitefish, *Coregonus lavaretus*. *J Anim Sci* **89:** 959-971.

Kruuk, L.E.B., Clutton-Brock, T.H., Slate, J., Pemberton, J.M., Brotherstone, S. & Guinness, F.E. 2000. Heritability of fitness in a wild mammal population. *Proc. Natl. Acad. Sci. USA* **97:** 698-703.

Kuparinen, A. & Hutchings, J.A. 2017. Genetic architecture of age at maturity can generate divergent and disruptive harvest-induced evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* **33:** 402-416.

Lande, R. 1982. A quantitative genetic theory of life-history evolution. *Ecology* **63:** 607-615.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37:** 1210-1226.

Lane, J.E., Kruuk, L.E., Charmantier, A., Murie, J.O., Coltman, D.W., Buoro, M., Raveh, S. & Dobson, F.S. 2011. A quantitative genetic analysis of hibernation emergence date in a wild population of Columbian ground squirrels. *J Evol Biol* **24:** 1949-1959.

Lee, S.H., Yang, J., Goddard, M.E., Visscher, P.M. & Wray, N.R. 2012. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28:** 2540-2542.

Mackay, T.F., Stone, E.A. & Ayroles, J.F. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.* **10:** 565-577.

Mackay, T.F.C. 2001. The genetic architecture of quantitative traits. *Annu. Rev. Genet.* **35:** 303-339.

Metcalfe, N.B., Van Leeuwen, T.E. & Killen, S.S. 2016. Does individual variation in metabolic phenotype predict fish behaviour and performance? J. Fish. Biol. **88**: 298-321.

Mills, K.E., Pershing, A.J., Sheehan, T.F. & Mountain, D. 2013. Climate and ecosystem linkages explain widespread declines in north american Atlantic salmon populations. *Glob. Chang. Biol.* **19:** 3046-3061.

Narum, S.R., Hatch, D., Talbot, A.J., Moran, P. & Powell, M.S. 2008. Iteroparity in complex mating systems of steelhead *Oncorhynchus mykiss* (Walbaum). *J. Fish. Biol.* **72:** 45-60.

Niemelä, E., Makinen, T.S., Moen, K., Hassinen, E., Erkinaro, J., Lansman, M. & Julkunen, M. 2000. Age, sex ratio and timing of the catch of kelts and ascending Atlantic salmon in the subarctic River Teno. *J. Fish. Biol.* **56:** 974-985.

Niemelä, E., Erkinaro, J., Julkunen, M., Hassinen, E., Lansman, M. & Brors, S. 2006a. Temporal variation in abundance, return rate and life histories of previously spawned Atlantic salmon in a large subarctic river. *J. Fish. Biol.* **68:** 1222-1240.

Niemelä, E., Orell, P., Erkinaro, J., Dempson, J.B., BrØrs, S., Svenning, M.A. & Hassinen, E. 2006b. Previously spawned Atlantic salmon ascend a large subarctic river earlier than their maiden counterparts. *J. Fish. Biol.* **69:** 1151-1163.

Nussey, D.H., Wilson, A.J., Morris, A., Pemberton, J., Clutton-Brock, T. & Kruuk, L.E. 2008. Testing for genetic trade-offs between early- and late-life reproduction in a wild red deer population. *Proc Biol Sci* **275:** 745-750.

Orr, H.A. 2000. Adaptation and the cost of complexity. *Evolution* **54:** 13-20.

Ozerov, M., Vasemagi, A., Wennevik, V., Niemela, E., Prusov, S., Kent, M. & Vaha, J.P. 2013. Cost-effective genome-wide estimation of allele frequencies from pooled DNA in Atlantic salmon (*Salmo salar* L.). *BMC Genomics* **14:** 12.

Paaby, A. B. & Rockman, M. V. 2013. The many faces of pleiotropy. *Trends. Genet.* **29**: 66-73.

Penney, Z.L. & Moffitt, C.M. 2013. Histological assessment of organs in sexually mature and post-spawning steelhead trout and insights into iteroparity. *Rev. Fish Biol. Fisher.* **24:** 781-801.

Piou, C. & Prévost, E. 2012. A demo-genetic individual-based model for Atlantic salmon populations: Model structure, parameterization and sensitivity. *Ecol. Model.* **231:** 37-52.

Piou, C. & Prevost, E. 2013. Contrasting effects of climate change in continental vs. oceanic environments on population persistence and microevolution of Atlantic salmon. *Glob. Chang. Biol.* **19:** 711-723.

Pritchard, V.L., Makinen, H., Vaha, J.P., Erkinaro, J., Orell, P. & Primmer, C.R. 2018. Genomic signatures of fine-scale local selection in atlantic salmon suggest involvement of sexual maturation, energy homeostasis, and immune defence-related genes. *Mol. Ecol*.

Reed, T.E., Schindler, D.E., Hague, M.J., Patterson, D.A., Meir, E., Waples, R.S. & Hinch, S.G. 2011. Time to evolve? Potential evolutionary responses of fraser river sockeye salmon to climate change and effects on persistence. *PLoS One* **6:** e20380.

Reid, J.E. & Chaput, G. 2012. Spawning history influence on fecundity, egg size, and egg survival of Atlantic salmon (*Salmo salar*) from the Miramichi River, New Brunswick, Canada. *Ices J. Mar. Sci.* **69:** 1678-1685.

Ritland, K. 1996. A Marker-based method for inferences about quantitative inheritance in natural populations. *Evolution* **50:** 1062-1073.

Robinson, M.R., Wilson, A.J., Pilkington, J.G., Clutton-Brock, T.H., Pemberton, J.M. & Kruuk, L.E. 2009. The impact of environmental heterogeneity on genetic architecture in a wild population of Soay sheep. *Genetics* **181:** 1639-1648.

Robinson, M.R., Santure, A.W., Decauwer, I., Sheldon, B.C. & Slate, J. 2013. Partitioning of genetic variation across the genome using multimarker methods in a wild bird population. *Mol. Ecol.* **22:** 3963-3980.

Rodriguez, J. A., Marigorta, U. M., Hughes, D. A., Spataro, N., Bosch, E. & Navarro, A. 2017. Antagonistic pleiotropy and mutation accumulation influence human senescence and disease. *Nat. Ecol. Evol.* **1**: 55.

Roff, D.A. 1996. The Evolution of Genetic Correlations: An Analysis of Patterns. *Evolution* **50:** 1392-1403.

Sanchez-Guillen, R.A., Wellenreuther, M. & Cordero Rivera, A. 2012. Strong asymmetry in the relative strengths of prezygotic and postzygotic barriers between two damselfly sister species. *Evolution* **66:** 690-707.

Santure, A.W., De Cauwer, I., Robinson, M.R., Poissant, J., Sheldon, B.C. & Slate, J. 2013. Genomic dissection of variation in clutch size and egg mass in a wild great tit (*Parus major*) population. *Mol. Ecol.* **22:** 3949-3962.

Schnurr, T.M., Gjesing, A.P., Sandholt, C.H., Jonsson, A., Mahendran, Y., Have, C.T., Ekstrom, C.T., Bjerregaard, A.L., Brage, S., Witte, D.R., Jorgensen, M.E., Aadahl, M., Thuesen, B.H., Linneberg, A., Eiberg, H., Pedersen, O., Grarup, N., Kilpelainen, T.O. & Hansen, T. 2016. Genetic Correlation between Body Fat Percentage and Cardiorespiratory Fitness Suggests Common Genetic Etiology. *PLoS One* **11:** e0166738.

Scott, G. R., Elogio, T. S., Lui, M. A., Storz, J. F. & Cheviron, Z. A. 2015. Adaptive modifications of muscle phenotype in high-altitude deer mice are associated with evolved changes in gene regulation. *Mol. Biol. Evol.* **32**: 1962-76.

Seamons, T.R. & Quinn, T.P. 2009. Sex-specific patterns of lifetime reproductive success in single and repeat breeding steelhead trout (*Oncorhynchus mykiss*). *Behav. Ecol. Sociobiol.* **64:** 505-513.

Sheldon, B.C., Kruuk, L.E. & Merila, J. 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. *Evolution* **57:** 406-420.

Simon, M.N., Machado, F.A. & Marroig, G. 2016. High evolutionary constraints limited adaptive responses to past climate changes in toad skulls. *Proc. Biol. Sci.* **283**.

Stearns, F. W. 2010. One hundred years of pleiotropy: a retrospective. *Genetics* **186**: 767-73.

Stearns, S., Dejong, G. & Newman, B. 1991. The effects of phenotypic plasticity on genetic correlations. *Trends Ecol. Evol.* **6:** 122-126.

745 Steppan, S.J., Phillips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the
746 G matrix. *Trends Ecol. Evol.* **17:** 320-327.
747 Storz, J.F., Bridgham, J.T., Kelly, S.A. & Garland, T., Jr. 2015. Genetic approaches in comparative
748 and evolutionary physiology. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **309:** R197-214.
749 Taranger, G.L., Carrillo, M., Schulz, R.W., Fontaine, P., Zanuy, S., Felip, A., Weltzien, F.A.,
750 Dufour, S., Karlsen, O., Norberg, B., Andersson, E. & Hansen, T. 2010. Control of puberty
751 in farmed fish. *Gen. Comp. Endocrinol.* **165:** 483-515.
752 Theriault, V., Garant, D., Bernatchez, L. & Dodson, J.J. 2007. Heritability of life-history tactics and
753 genetic correlation with body size in a natural population of brook charr (*Salvelinus*
754 *fontinalis*). *J. Evol. Biol.* **20:** 2266-2277.
755 Thorstad, E.B., Whoriskey, F., Rikardsen, A.H. & Aarestrup, K. 2010. Aquatic Nomads: The Life
756 and Migrations of the Atlantic Salmon. In: *Atlantic Salmon Ecology*, pp. 1-32. Wiley-
757 Blackwell.
758 Vähä, J.-P., Erkinaro, J., Falkegård, M., Orell, P. & Niemelä, E. 2017. Genetic stock identification
759 of Atlantic salmon and its evaluation in a large population complex. *Can. J. Fish. Aquat.*
760 *Sci.* **74:** 327-338.
761 Vähä, J.P., Erkinaro, J., Niemelä, E. & Primmer, C.R. 2007. Life-history and habitat features
762 influence the within-river genetic structure of Atlantic salmon. *Mol. Ecol.* **16:** 2638-2654.
763 Wagner, G.P. 1989. Multivariate mutation-selection balance with constrained pleiotropic effects.
764 *Genetics* **122:** 223-234.
765 Wagner, G.P. & Zhang, J. 2011. The pleiotropic structure of the genotype-phenotype map: the
766 evolvability of complex organisms. *Nat. Rev. Genet.* **12:** 204-213.
767 Walters, C. & Maguire, J.J. 1996. Lessons for stock assessment from the northern cod collapse.
768 *Rev. Fish Biol. Fisher.* **6:** 125-137.
769 Wang, Z., Liao, B.Y. & Zhang, J. 2010. Genomic patterns of pleiotropy and the evolution of
770 complexity. *Proc. Natl. Acad. Sci. USA* **107:** 18034-18039.
771 Weigel, K.A., VanRaden, P.M., Norman, H.D. & Grosu, H. 2017. A 100-Year Review: Methods
772 and impact of genetic selection in dairy cattle-From daughter-dam comparisons to deep
773 learning algorithms. *J. Dairy. Sci.* **100:** 10234-10250.
774 Wilson, A.J., Reale, D., Clements, M.N., Morrissey, M.M., Postma, E., Walling, C.A., Kruuk, L.E.
775 & Nussey, D.H. 2010. An ecologist's guide to the animal model. *J. Anim. Ecol.* **79:** 13-26.
776 Yang, J., Benyamin, B., McEvoy, B.P., Gordon, S., Henders, A.K., Nyholt, D.R., Madden, P.A.,
777 Heath, A.C., Martin, N.G., Montgomery, G.W., Goddard, M.E. & Visscher, P.M. 2010.
778 Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.*
779 **42:** 565-569.
780

Main text tables:

Table 1: Coefficients of the most parsimonious model for survival to second spawning (repeat spawning) in Atlantic salmon at the *vgll3*$_{\text{Met54Thr}}$ locus. Sea age denotes the sea age at first maturity. *E* and *L* denote alleles linked to the early and late age at maturity, respectively. See Table S5 for a full list of all plausible models.

Coefficient	Estimate	Std.Error	*z*-value	*p*-value
Inarijoki	0.656	0.452	1.45	0.147
Tenojoki	0.682	0.523	1.30	0.192
Sex (Male)	-0.008	0.243	-0.03	0.973
Sea age	-0.686	0.194	-3.53	0.000
SNP (*EL*)	-0.786	0.508	-1.55	0.122
SNP (*EE*)	0.050	0.621	0.08	0.936
Sea Age : SNP (*EL*)	0.706	0.228	3.10	0.002
Sea Age : SNP (*EE*)	0.323	0.347	0.93	0.351
Tenojoki : Sex (Male)	-0.837	0.346	-2.42	0.015

Table 2: Coefficients of the most parsimonious model explaining variation in total length at capture in first-time spawner and repeat spawner groups. The first-time spawner model includes *vgll3*$_{Met54Thr}$, in which alleles are coded additively (i.e., *EE*=1, *EL*=2, and *LL* = 3). Length is log-normalized. See Table S10 for a list of all plausible models.

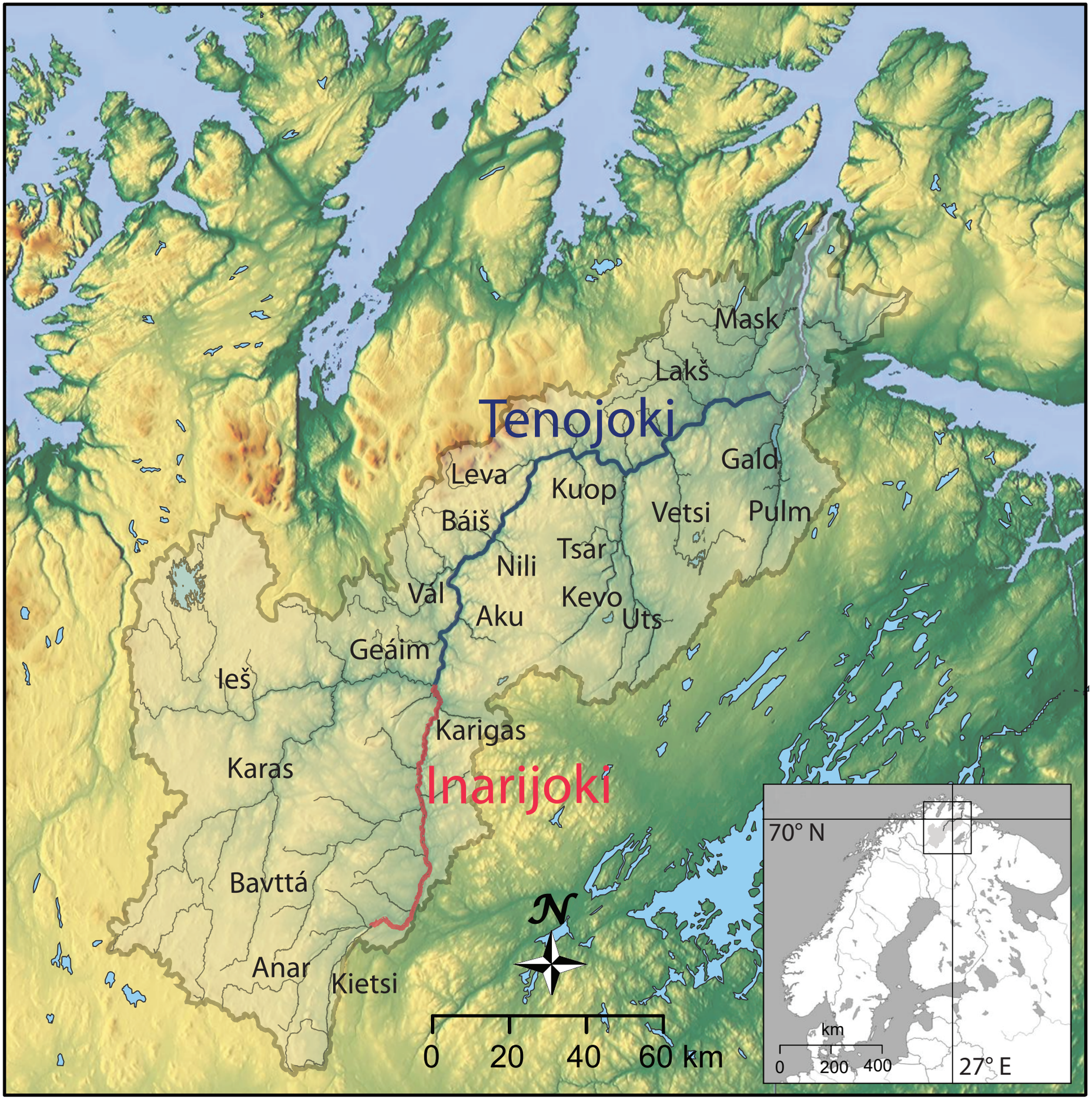
	Coefficient	Estimate	Std. Error	*t*- value	*p*-value
First-time spawners	Intercept	3.755	0.022	171.80	<0.001
	SNP	0.021	0.005	3.90	<0.001
	Population	0.223	0.044	5.09	<0.001
	Sea age	0.258	0.011	22.83	<0.001
	Sex	-0.031	0.032	-0.98	0.329
	Population : Sea age	-0.072	0.018	-4.06	<0.001
	Population : Sex	-0.185	0.052	-3.56	<0.001
	Sea age : Sex	0.053	0.022	2.39	0.017
	Population : Sea age : Sex	0.056	0.027	2.10	0.037
Repeat spawners	Intercept	4.226	0.016	267.37	<0.001
	Population	0.080	0.011	7.19	<0.001
	Sea age	0.139	0.008	16.41	<0.001
	Sex	0.095	0.023	4.18	<0.001
	Sea age : Sex	-0.049	0.012	-4.00	<0.001

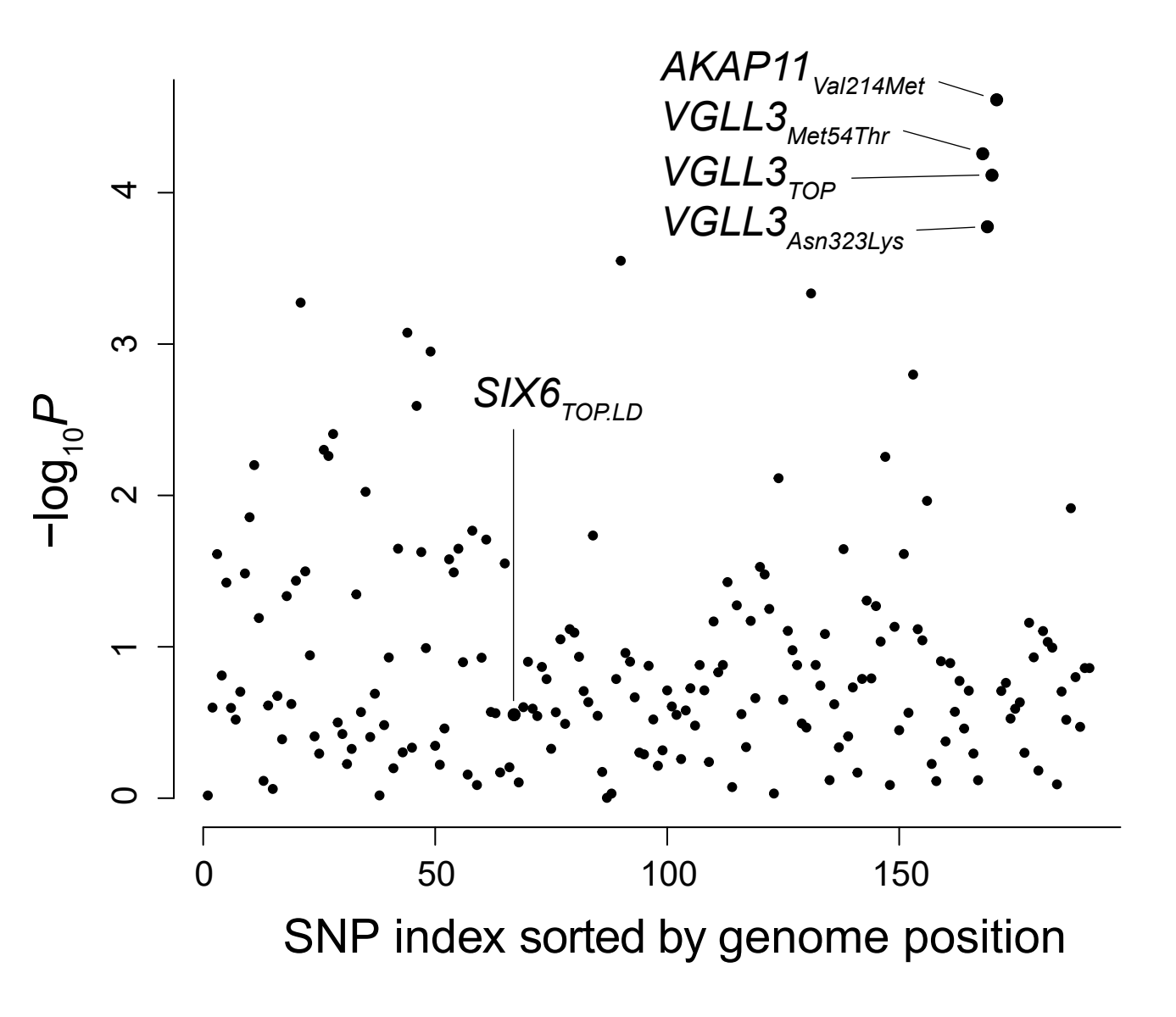
Figure legends:

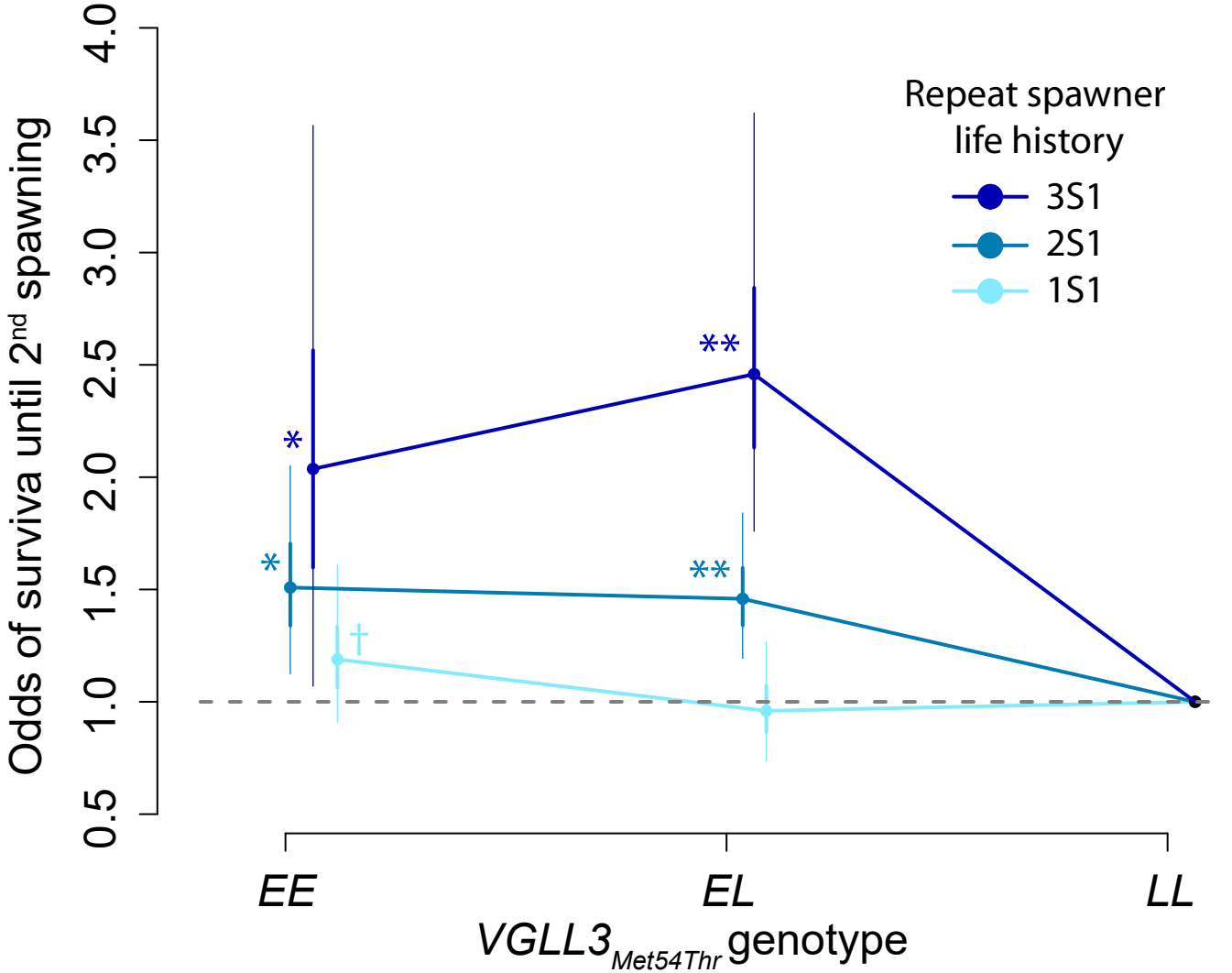
Figure 1: Map of the study system. River sections with blue and red colours indicate Tenojoki and Inarijoki, respectively. See Table S2 for full river names.

Figure 2: Goodness of the fit of the genetic models in 188 SNPs. *P*-values indicate the significance of the fit of the genetic model relative to the null model. SNPs previously associated with age at maturity are labelled.

Figure 3: Odds of survival to second spawning (repeat spawning) of the *EE* and *EL* genotypes, relative to the *LL* genotype, at the $vgll3_{Met54Thr}$ locus, as estimated by 10000 parametric permutations. Dots, thick lines, and thin lines denote median estimates, 50% CIs, and 90% CIs, respectively. Asterisks denote significance, as calculated by the proportion of permutations whose odds of survival were greater than that of the *LL* genotype (* $p<0.05$, ** $p<0.001$; † denotes the proportion of permutations for which the odds of survival were greater than that of the *EL* genotype at $p=0.051$). See Table 1 for full model specifications.


Mask
Lakš
Tenojoki
Gald
Leva
Kuop
Vetsi
Pulm
Báiš
Tsar
Nili
Kevo
Väl
Aku
Uts
Geáim
Ieš
Karigas
Karas
Inarijoki
Bavttá
Anar
Kietsi
N
0 20 40 60 km
70° N
km
0 200 400
27° E


AKAP11 Val214Met
VGLL3 Met54Thr
VGLL3 TOP
VGLL3 Asn323Lys
SIX6 TOP.LD
−log10P
0
1
2
3
4
0
50
100
150
SNP index sorted by genome position


Odds of surviva until 2nd spawning
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Repeat spawner life history
3S1
2S1
1S1
*
*
**
**
†
EE
EL
LL
VGLL3Met54Thr genotype

Supporting Information:

Table S1: Sampling month and year of repeat and first-time spawners used in the study.

	Year of sampling	May	June	July	August
Repeat spawners	2001	0	9	27	4
	2002	6	32	13	16
	2003	3	22	34	21
	2004	0	37	14	8
	2005	0	7	14	20
	2006	0	7	7	5
	2007	0	13	24	9
	2008	0	18	49	18
	2009-14	0	20	13	8
First-time spawners	2001	0	0	0	221
	2002	0	0	0	213
	2003	0	0	0	209

Table S2: Assignment of individuals to origin of populations. An individual's likelihood to originate from a population was estimated by comparing it to the self-assignment likelihood distribution of 1000 simulated individuals in every baseline population*. An individual was assigned to a baseline population if its likelihood was not less than 1% of the likelihood distribution of the simulated data, and if the likelihood score of the individual to the primary population candidate was at least 10 times greater than the likelihood to the second best candidate.

Figure 1 abbreviations†	Baseline Populations[2]	First-time spawners	Repeat spawners	Juveniles‡	
				MS	Inari
	Tenojoki	335	115	143	0
	Inarijoki	167	141	0	98
Ieš	Iešjohka (33)	8	17	0	1
Karas	Kárášjohka (52)	10	12	0	0
Bavttá	Bavttájohka (50)	0	0	0	0
Geáim	Geáimmejohka (45)	1	2	0	0
Kietsi	Kietsimäjoki (72)	1	0	0	0
Anar	Anarjohka (46)	0	0	0	0
Karigas	Karigasjoki (39)	2	3	0	0
Vál	Váljohka (45)	1	11	0	0
Aku	Akujoki (40)	6	4	0	0
Báiš	Báišjohka (36)	1	2	0	0
Nili	Nilijoki (42)	1	2	0	0
Leva	Levajohka (54)	0	1	0	0
Kuop	Kuoppilasjoki (55)	0	4	0	0
Tsar	Tsarsjoki (48)	0	10	0	0
Kevo	Kevojoki (80)	0	8	0	0
Uts	Utsjoki (52)	0	1	0	0
Vetsi	Vetsijoki (27)	3	7	0	0
Lakš	Lakšjohka (39)	0	1	0	0
Pulm	Ylä-Pulmankijoki (32)	0	0	0	0
Gald	Galddasjoki (70)	0	0	0	0
Mask	Máskejohka (70)	0	0	0	0
no assignment (type 1)§		8	6	1	0
no assignment (type 2)¶		99	131	16	21
% assignment to Tenojoki and Inarijoki		78.1	53.6	89.4	81.7
TOTAL		643	478	160	120

* Baseline population minor allele frequencies were adjusted to a minimum of 0.05 to avoid very small to zero probabilities, and likelihood distributions were always calculated after accounting for the missing SNPs for any candidate individual.

† Baseline population SNP data information were extracted from obtained by genotype by pooling (Ozerov *et al.*, 2013). Number of individuals in each pool (10 ng/ul DNA from every individuals) is given in parenthesis. (Minimum 2 replicates per population. See Ozerov *et al.* (2013) and Vähä *et al.* (2017) for details.) For Tenojoki and Inarijoki, baseline data were generated by performing individual genotyping and by combining data from juvenile fish (collected by electrofishing within populations' boundaries), and adult migrating fish that were previously assigned to these populations ($N_{Tenojoki}$ =347 and $N_{Inarijoki}$ =171) (Aykanat *et al.*, 2015).

‡ Juvenile fish sampled within Tenojoki and Inarijoki population range.

§ Individual that has low likelihood for all populations.

¶ Individual that has high likelihood of origin for more than one population.

Table S3: Individuals assigned to population of origin as a function of life history, sex and capture time in the season. All maiden fish were collected in August, while repeat spawner fish were collected from May-August (see also Table S1).

	Population of origin*	First-time spawners			Repeat spawners					
		August			May-July			August		
		1SW	2SW	3SW	1S1	2S1	3S1	1S1	2S1	3S1
Male	Inari	60	18	2	45	13	1	7	2	2
	Tenojoki	66	28	84	8	12	4	7	11	2
	Others	49	10	11	43	22	7	7	2	0
	TOTAL	175	56	97	96	47	12	21	15	4
Female	Inari	35	42	10	28	14	2	18	9	0
	Tenojoki	2	38	117	3	10	39	1	5	13
	Others	16	34	21	45	43	30	7	10	6
	TOTAL	53	114	148	76	67	71	26	24	19

* Population of origin as inferred in Table S2. Others denote individuals that were assigned to non-focal populations or unassigned.

Table S4: All possible null model structures (i.e. without the genetic effects) explaining repeat spawner prevalence in Tenojoki and Inarijoki populations. Plus (+) sign and a numeric value in coefficient columns indicates that the given factor was included in the model as a categorical, or a numeric variable, respectively. Empty cell indicates that the factor was not specified in the model.

Model coefficients (repeat spawner prevalence)								df	logLik	AICc	delta	weight
intercept	*pop*	*sex*	*sea age*	*pop* : *sex*	*pop* : *sea age*	*sex* : *sea age*	*pop* : *sex* : *sea age*					
0.42	+	+	-0.40	+				5	-459.12	928.33	0.00	0.261
1.18	+	+	-0.90	+	+	+	+	8	-456.12	928.43	0.10	0.248
0.72	+	+	-0.59	+		+		6	-458.23	928.58	0.25	0.231
0.55	+	+	-0.48	+	+			6	-458.99	930.09	1.77	0.108
0.75	+	+	-0.62	+	+	+		7	-458.22	930.59	2.26	0.084
0.85	+	+	-0.57		+			5	-461.67	933.43	5.10	0.020
0.53	+	+	-0.33					4	-462.70	933.46	5.13	0.020
0.74	+	+	-0.51		+	+		6	-461.54	935.18	6.86	0.008
0.43	+	+	-0.29			+		5	-462.60	935.27	6.94	0.008
-0.20	+	+		+				4	-464.29	936.63	8.30	0.004
0.50	+		-0.47		+			4	-465.03	938.11	9.78	0.002
0.13	+		-0.21					3	-466.29	938.62	10.29	0.002
-0.03	+	+						3	-466.54	939.12	10.79	0.001
-0.17	+							2	-468.15	940.32	11.99	0.001
0.66		+	-0.54					3	-467.51	941.04	12.72	0.000
0.68		+	-0.55			+		4	-467.50	943.05	14.73	0.000
0.16			-0.43					2	-473.51	951.03	22.71	0.000
-0.54		+						2	-483.16	970.33	42.00	0.000
-0.67								1	-484.76	971.52	43.19	0.000

Table S5: Genetic models* explaining repeat spawner prevalence in the Tenojoki and Inarijoki populations at four chromosome 25 SNPs associated with sea age at maturity. Only more likely models are listed for each SNP (i.e. ΔAICc<2 to the most parsimonious model within a SNP). Plus (+) sign and a numeric value in coefficient columns indicates that the given factor was included in the model as a categorical, or a numeric variable, respectively†. Empty cell indicates that the factor was not specified in that particular model. Only samples successfully genotyped for all four loci were included in the analysis, enabling direct comparison of AICc values between models with that parameterise different SNPs (i.e. ΔAICcall column).

SNP name	Model coefficients (repeat spawner prevalence)										df	logLik	AICc	ΔAICc	ΔAICc all
	Intercept	*SNP*†	*pop*	*sex*	*sea age*	*SNP : pop*	*SNP : sex*	*SNP[1] : sea age*	*pop : sex*	*SNP:pop:sex*					
vgll3$_{TOP}$	-0.344	0.270	+	+	0.829			-0.450	+		7	-371.7	757.63	0.00	-3.12
	-0.514	0.310	+	+	1.187	+		-0.617	+		8	-370.8	757.78	0.15	-3.27
	-1.037	0.613	+	+	1.350	+	+	-0.694	+		9	-370.0	758.23	0.60	-3.72
	-0.803	0.540	+	+	0.950		+	-0.508	+		8	-371.1	758.34	0.72	-3.84
	-0.921	0.487	+	+	1.490	+	+	-0.744	+	+	10	-369.4	759.16	1.54	-4.66
	-0.047	+	+	+	0.433			+	+		9	-370.7	759.75	2.12	-5.24
vgll3$_{Met54Thr}$	0.930	+	+	+	-0.256			+	+		9	-368.1	754.51	0.00	0.00
	0.922	+	+	+	-0.233		+	+	+		11	-367.0	756.53	2.02	-2.02
	0.072	0.108	+	+	0.682			-0.376	+		7	-371.2	756.69	2.18	-2.18
vgll3$_{Asn323Lys}$	0.971	+	+	+	-0.146		+	+	+		11	-369.3	761.09	0.00	-6.58
	0.915	+	+	+	-0.174			+	+		9	-372.1	762.41	1.32	-7.90
	-0.092	0.167	+	+	0.760			-0.386	+		7	-374.4	762.90	1.81	-8.39
	0.609	+	+	+	0.262	+	+	+	+		13	-368.4	763.32	2.23	-8.81
akap11$_{Val214Met}$	-1.627	0.760	+	+	-0.221				+		6	-373.0	758.11	0.00	-3.60
	-1.377	0.666	+	+	-0.213		+		+		7	-372.8	759.74	1.63	-5.23
	-0.942	0.507	+	+	-0.524			0.115	+		7	-372.8	759.83	1.72	-5.32
	-1.051	+	+	+	-0.220				+		7	-372.9	759.98	1.87	-5.47
	-1.368	0.665	+	+	-0.217	+			+		7	-372.9	760.04	1.93	-5.53
	-0.171	0.221	+	+	-0.691		+	0.181	+		8	-372.4	761.09	2.98	-6.58
Optimum Null	-0.648		+	+					+		5	-383.1	776.28		-21.77

* Genetic effects were modelled onto the best null model structure as inferred in the Table S2. †A numeric value or a plus sign in the SNP coefficient indicates that the SNP effect was coded additively (i.e. allelic effect with df=1), or non-additively (genotype effect with df=2), respectively. For additive models, genotypes are called as *EE*=1, *EL*=2, and *LL*=3.

Table S6: Number of individuals per genotype in *vgll3*$_{\text{Met54Thr}}$ sampled within each population and life history type. Female and male numbers in each group given in parentheses.

Populations	*Genotype**	Life history (numbers)					
		1SW	1S1	2SW	2S1	3SW	3S1
Tenojoki	*LL*	19 (0,19)	7 (2,5)	26 (10,16)	13 (7,6)	129 (74,55)	22 (19,3)
	EL	33 (1,32)	10 (1,9)	18 (11,7)	19 (5,14)	45 (28,17)	30 (27,3)
	EE	5 (0,5)	2 (1,1)	4 (3,1)	4 (1,3)	5 (4,1)	4 (4,0)
Inarijoki	*LL*	15 (3,12)	14 (5,9)	14 (11,3)	6 (4,2)	7 (7,0)	3 (1,2)
	EL	44 (18,26)	39 (16,23)	27 (16,11)	25 (16,9)	2 (2,0)	2 (1,1)
	EE	29 (11,18)	43 (23,20)	9 (8,1)	7 (3,4)	1 (1,0)	0 (0,0)

* Alleles are named reflecting their association with early (*E*) or late (*L*) age at maturation at sea.

Table S7: Predicted odds of survival to second spawning as a function of genotype at *vgll3*$_{\text{Met54Thr}}$, within the given sea age at first maturity, population and sex. Quantiles of odds estimation and P_{relLL} are after 10000 parametric permutations and relative to *LL* genotype. P_{relEL} is the significance of difference in odds of survival to second spawning in *EE* genotype relative to *EL* genotype. P_{Add} is the significance of models probability to deviate from additivity.

vgll3$_{Met54Thr}$	Sea age	Sex	Pop	Quantiles after parametric permutations					$P_{\text{rel}LL}$	$P_{\text{rel}EL}$	P_{Add}
				5	25	50	75	95			
EL	1	M	Ina	0.74	0.87	0.96	1.07	1.26	0.600	0.053	0.138
EE	1	M	Ina	0.91	1.06	1.18	1.32	1.57	0.144		
EL	2	M	Ina	1.19	1.33	1.44	1.58	1.79	0.001	0.405	0.081
EE	2	M	Ina	1.12	1.33	1.49	1.67	1.98	0.012		
EL	3	M	Ina	1.74	2.10	2.42	2.77	3.41	0.000	0.713	0.014
EE	3	M	Ina	1.08	1.59	2.02	2.51	3.40	0.033		
EL	1	F	Ina	0.67	0.82	0.94	1.10	1.37	0.600	0.053	0.126
EE	1	F	Ina	0.88	1.09	1.27	1.48	1.88	0.144		
EL	2	F	Ina	1.26	1.46	1.61	1.79	2.09	0.001	0.405	0.115
EE	2	F	Ina	1.15	1.46	1.69	1.96	2.45	0.012		
EL	3	F	Ina	2.01	2.52	2.94	3.47	4.38	0.000	0.714	0.030
EE	3	F	Ina	1.10	1.74	2.34	3.10	4.54	0.033		
EL	1	M	Teno	0.74	0.86	0.96	1.07	1.27	0.600	0.053	0.136
EE	1	M	Teno	0.91	1.06	1.19	1.33	1.59	0.144		
EL	2	M	Teno	1.19	1.34	1.45	1.59	1.80	0.001	0.405	0.084
EE	2	M	Teno	1.12	1.34	1.50	1.69	1.99	0.012		
EL	3	M	Teno	1.76	2.12	2.42	2.79	3.50	0.000	0.714	0.015
EE	3	M	Teno	1.08	1.60	2.02	2.54	3.48	0.033		
EL	1	F	Teno	0.75	0.86	0.96	1.07	1.27	0.600	0.053	0.135
EE	1	F	Teno	0.91	1.06	1.18	1.33	1.59	0.144		
EL	2	F	Teno	1.19	1.34	1.45	1.58	1.80	0.001	0. 405	0.083
EE	2	F	Teno	1.12	1.34	1.50	1.69	2.00	0.012		
EL	3	F	Teno	1.76	2.12	2.42	2.79	3.45	0.000	0. 714	0.015
EE	3	F	Teno	1.08	1.60	2.03	2.55	3.47	0.033		

Table S8: Model coefficients and fit of the genetic models with different dominance architecture in the *vgll3*$_{Met54Thr}$ SNP.

Model		*popIna*	*popTeno*	*sex*	*sea age*	*SNP2**	*SNP3**	*sea age: SNP2*	*sea age: SNP3*	*popTeno: sex*	AIC	ΔAIC	Description
1	Est.	0.706	-0.008	-0.836	0.383	-0.837	0.580	0.313	0.621	0.347	843.45	5.62	Partial dominance allowed. (Genotypes coded as factors.)
	S.E.	0.732	-0.362	-0.050	-0.323	0.488	0.243	0.546	0.333	0.346			
2	Est.	0.679	0.016	-0.209	-0.883	0.519		0.192		0.207	837.83	0.00	Age dependent genetic architecture. (complete dominance for *L* allele in 1S1, and *E* allele in 2S1 and 3S1 life histories)
	S.E.	0.789	-0.318	-0.325	0.387	0.244		0.437		0.347			
3	Est.	0.181	-0.018	0.573	-0.897	0.479		0.286		0.311	862.64	24.81	Age dependent genetic architecture. (complete dominance for *E* allele in 1S1, and *L* allele in 2S1 and 3S1 life histories)
	S.E.	0.134	-0.127	-0.383	0.392	0.241		0.503		0.341			
4	Est.	-0.125	0.008	0.171	-0.862	0.731		0.362		0.148	845.56	7.73	Additive SNP effect.
	S.E.	-0.085	0.505	-0.352	0.649	0.243		0.293		0.345			
5	Est.	0.244	0.018	0.066	-0.866	0.733		0.365		0.138	843.46	5.64	Complete dominance for *L* allele.
	S.E.	0.322	0.348	-0.311	0.645	0.244		0.245		0.345			
6	Est.	-0.713	-0.001	0.486	-0.875	0.590		0.325		0.137	846.65	8.82	Complete dominance for *E* allele.
	S.E.	-0.710	0.733	-0.467	0.511	0.242		0.251		0.345			
7	Est.	0.534	0.006	-0.068	-0.873	0.689		0.276		0.119	839.89	2.06	Age dependent genetic architecture. (additive in 1S1, and dominant for *E* allele in 2S1 and 3S1 life histories)
	S.E.	0.609	-0.048	-0.178	0.579	0.243		0.274		0.346			

* vgll3$_{Met54Thr}$ *is modeled as a categorical factor in model 1, and additively in other models (heterozygotes coded identical to the homozygote genotype of the dominant allele). Genotypes are called as EE=1, EL=2, and LL=3.*

Table S9: Relation between sampling time in the season and genotype in the repeat spawner individuals (N=256).

Day of sampling (Julian calendar day)	Estimate	Std. Error	*t*-value	*P*
Pop Inarijoki	221.96	7.57	29.31	<0.001
Pop Tenojoki	235.24	9.79	24.04	<0.001
Age at first spawning	-5.56	2.21	-2.51	0.013
$vgll3_{Met54Thr}$ (*EE*=1, *EL*=2, *LL*=3)	-4.16	2.15	-1.94	0.054

Table S10: Most parsimonious coefficients explaining variation in total length at capture in first-time and repeat spawner groups, ascending from most parsimonious to less at *vgll3*$_{\text{Met54Thr}}$ locus. Forth order interactions were not parametrised by any most likely models, hence excluded from the table. Only models within five AICc difference to the most optimal model are shown. Null models (i.e. models without SNP as coefficients) are marked as NULL in the "model column". Plus (+) sign and a numeric value in coefficient columns indicates that the given factor was included in the model as a categorical, or a numeric variable, respectively[1]. Empty cell indicates that the factor was not specified in the model. (SA= sea age)

	Model coefficients (Length at capture)														df	logLik	AICc	ΔAICc	model
	Intercept	*SNP**	*pop*	*SA*	*sex*	*SNP: pop*	*SNP: SA*	*SNP: sex*	*pop: SA*	*pop: sex*	*SA: sex*	*SNP: pop:SA*	*SNP: pop:sex*	*SNP: SA:sex*	*pop: SA:sex*					
repeat	4.22		+	0.139	+						+					6	317.13	-621.91	0.00	null
spawners	4.22		+	0.142	+					+	+					7	317.30	-620.13	1.78	null
	4.23		+	0.135	+				+		+					7	317.23	-620.00	1.92	null
	4.22	0.001	+	0.139	+						+					7	317.15	-619.83	2.08	
	4.23	+	+	0.140	+						+					8	317.79	-618.99	2.93	
	4.22	0.007	+	0.138	+	+					+					8	317.76	-618.92	2.99	
	4.25	-0.011	+	0.106	+	+	0.014				+					9	318.70	-618.65	3.26	
first	3.76	0.021	+	0.258	+				+	+	+				+	10	563.37	-1106.22	0.00	
time	3.78	+	+	0.258	+				+	+	+				+	11	564.22	-1105.80	0.42	
spawners	3.76	0.014	+	0.261	+			+	+	+	+				+	11	564.12	-1105.61	0.61	
	3.76	0.015	+	0.261	+	+			+	+	+				+	11	563.82	-1105.00	1.22	
	3.77	0.009	+	0.263	+	+		+	+	+	+				+	12	564.55	-1104.34	1.88	
	3.76	0.020	+	0.258	+		0.000		+	+	+				+	11	563.37	-1104.11	2.11	
	3.80	+	+	0.249	+				+	+	+					10	562.27	-1104.01	2.20	
	3.78	0.004	+	0.252	+		0.004	+	+	+	+				+	12	564.34	-1103.92	2.29	
	3.77	0.021	+	0.249	+				+	+	+					9	561.14	-1103.85	2.37	
	3.80	+	+	0.249	+		+		+	+	+				+	13	565.15	-1103.41	2.81	
	3.78	0.014	+	0.251	+			+	+	+	+					10	561.87	-1103.21	3.01	

* A numeric value or a plus sign in the SNP coefficient indicates that the SNP effect was coded additively (i.e. allelic effect with df=1), or non-additively (genotype effect with df=2), respectively. For additive models, genotypes are called as *EE*=1, *EL*=2, and *LL*=3.

Table S11: Coefficients of most parsimonious models explaining post spawning growth among repeat spawner fish*. Only models within two AICc difference to the most optimal model are shown. Null models (i.e. models without SNP as coefficients) are marked as NULL in the "model column." Plus (+) sign and a numeric value in coefficient columns indicates that the given factor was included in the model as a categorical, or a numeric variable, respectively1. Empty cell indicates that the factor was not specified in the model. (SA= sea age)

Model coefficients (Post spawning growth)								df	logLik	AICc	ΔAICc	model	weight
(Intercept)	*SNP*	*pop*	*sea age*	*SNP:pop*	*SNP:SA*	*pop:SA*	*SNP:pop:SA*						
1.379			-0.294					3	7.29	-8.47	0.00	null	0.112
1.156	0.076	+	-0.149	+	-0.053	+		8	12.44	-8.13	0.33		0.095
1.329	0.032		-0.302					4	8.16	-8.12	0.35		0.094
1.190	0.100		-0.214		-0.041			5	9.15	-7.98	0.48		0.088
1.371		+	-0.299					4	7.94	-7.68	0.79	null	0.076
1.322		+	-0.269			+		5	8.96	-7.60	0.86	null	0.073
1.322	0.031	+	-0.307					5	8.77	-7.24	1.23		0.060
1.346	+		-0.305					5	8.72	-7.13	1.34		0.058
1.327	-0.003	+	-0.268	+		+		7	10.84	-7.10	1.37		0.057
1.252	0.032	+	-0.215	+	-0.023	+	+	9	12.94	-6.94	1.53		0.052
1.189	0.096	+	-0.222		-0.040			6	9.68	-6.93	1.53		0.052
1.280	0.029	+	-0.278			+		6	9.67	-6.90	1.57		0.051
1.136	0.098	+	-0.185		-0.042	+		7	10.71	-6.83	1.64		0.050
1.218	0.081	+	-0.207	+	-0.046			7	10.59	-6.59	1.87		0.044
1.364	0.010	+	-0.307	+				6	9.36	-6.29	2.18		0.038

* Post spawning growth is inferred from scale growth after the first spawning event. For additive models, genotypes are called as *EE*=1, *EL*=2, and *LL*=3.

† weight XXX andx YYY…

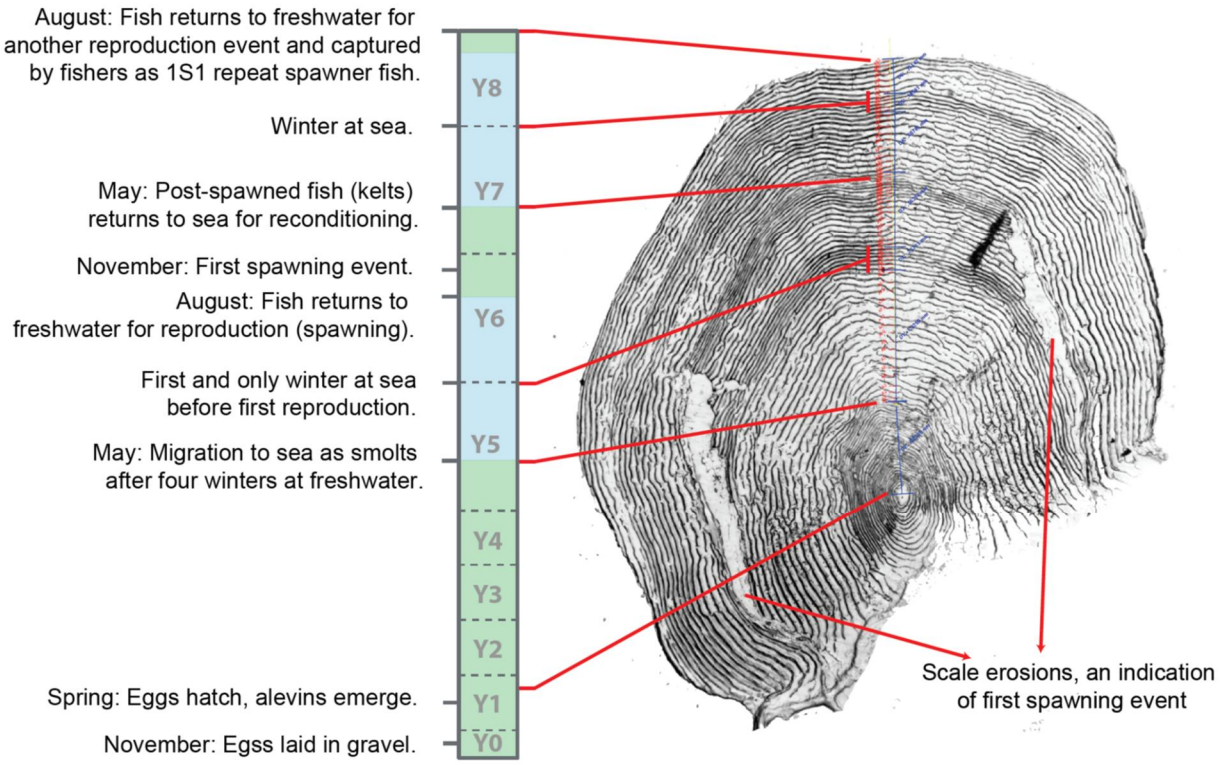


Figure S1: Life history diagram of a typical 1S1 repeat spawner fish with important event overlaid on the scale. Major events inferred from scale growth patterns are paired with red lines between the timeline column and the scale image. The timeline column also indicates freshwater and sea periods with green and blue, respectively, and year breaks are indicated with dashed lines. Note that the scale growth at a time, and the spacing between dark rings (circuli) is correlated with growth rate at that particular time point (see also Figure S3 and Figure S4).

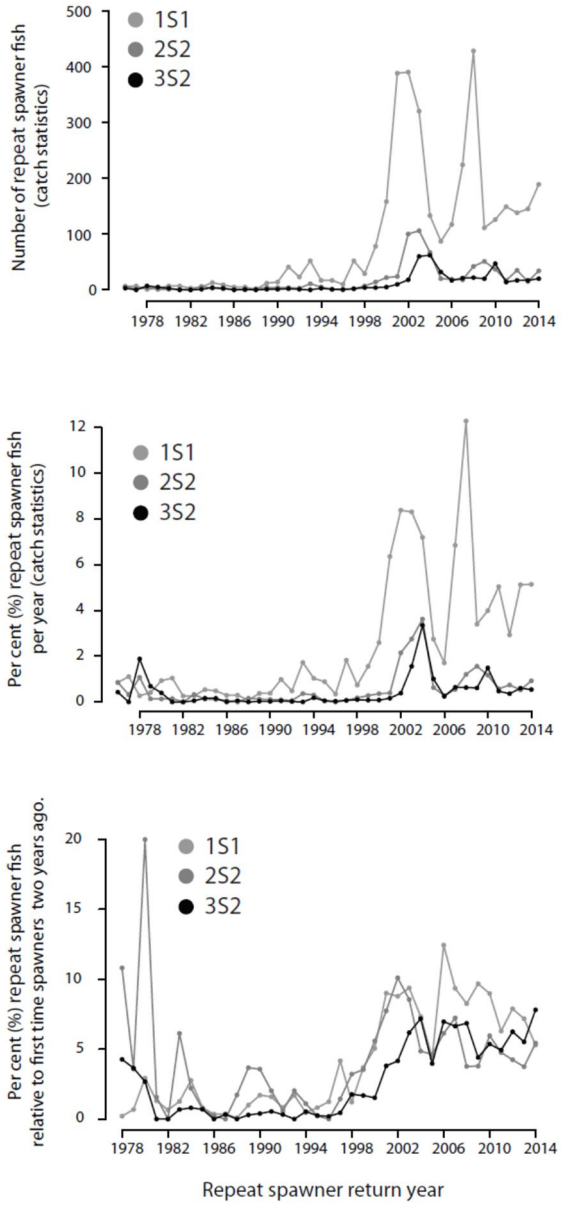


Figure S2: Changes in repeat spawner proportions in the Teno River over the last 40 years as inferred by absolute numbers (upper panel), relative to first-time spawner numbers in the same year (middle panel), or relative to the number of corresponding sea age class, two year earlier (lower panel). Note that the proportions are from mixed stock river fisheries data from scale collection based on Finnish catch statistics and do not reflect population specific proportions.

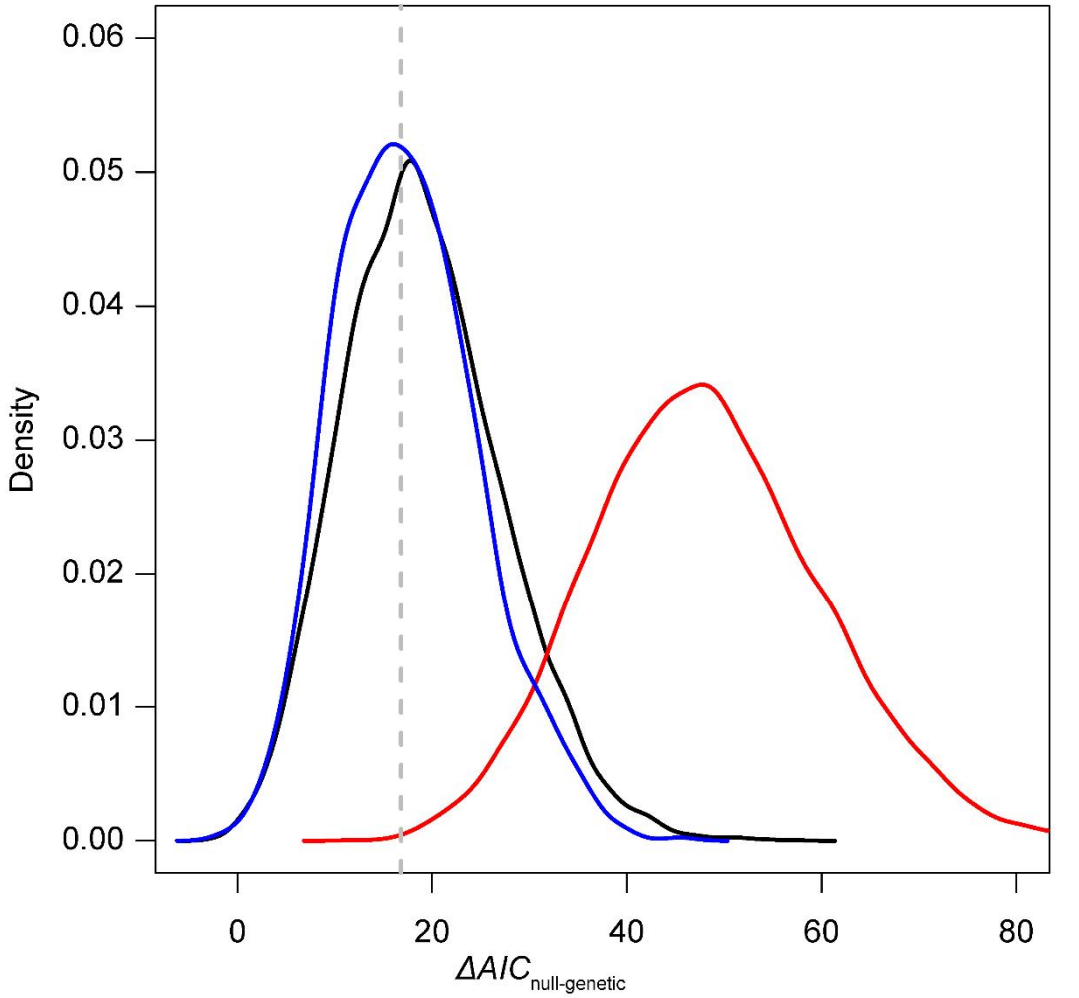
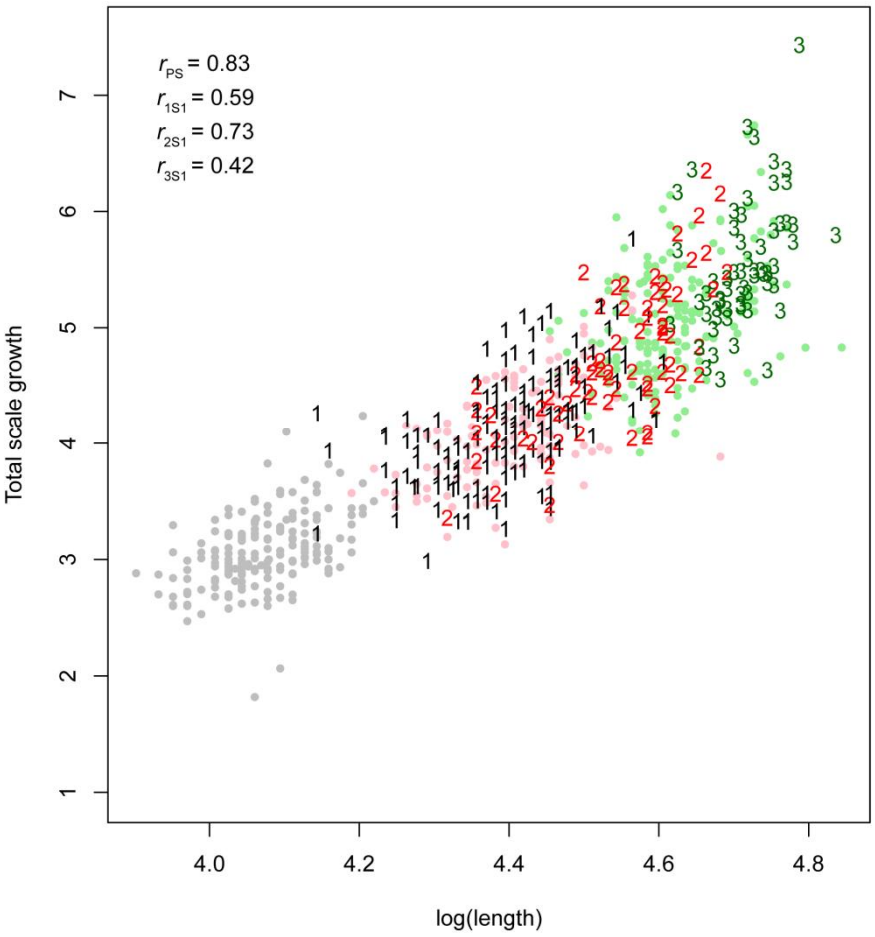


Figure S3: The sensitivity of sampling period of repeat spawner to the model fit ($\Delta AIC_{null\text{-}genetic}$). Black line shows the distribution of model fit of 10000 permuted models after repeat spawner fish were resampled (with replacement). In blue, same as above but resampled fish was constrained to match the average returning time of all repeat spawner fish captured in the Teno River. In red, same as above but resampled fish included only repeat-spawner fish that were captured in August (i.e. similar to sampling period for first-time spawners).



90

91 **Figure S4:** Correlation between total scale growth (mm) and log(length) in adult fish
92 sampled for this study. Numbers indicate the number of years spent at sea before first
93 spawning (i.e. 1 for 1S1, 2 for 2S1, and 3 for 3S1). Gray, pink and green dots shows the
94 relation for first-time spawners, for ages one to three years at sea, respectively. Pearson
95 correlation is given for each of the repeat spawner group combined and separately.

96

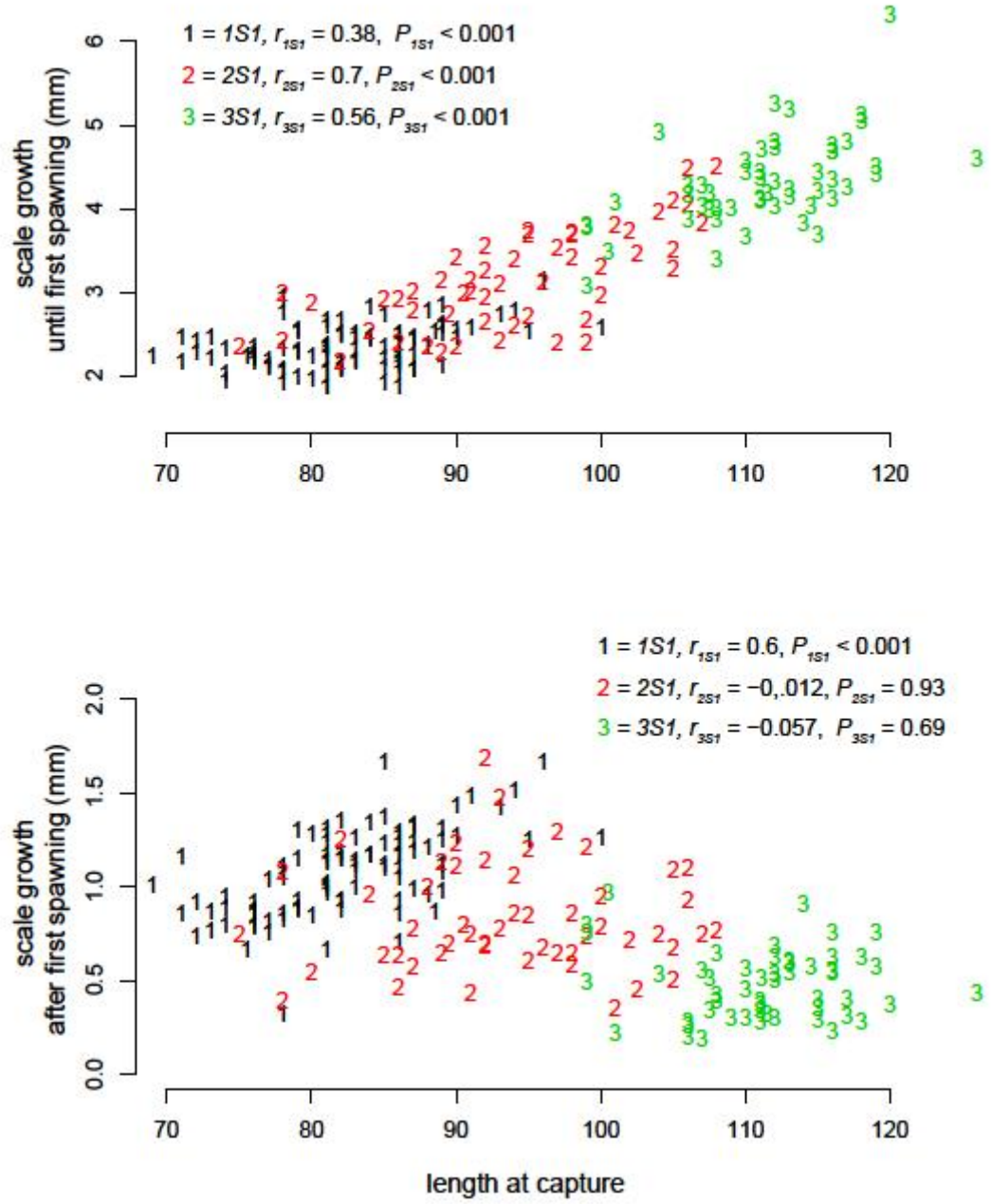


Figure S5: Correlation between length at capture and stage specific scale growth for 1S1, 2S1 and 3S1 repeat spawners, as given by scale growth before and after first spawning.

Appendix S1: Methods on scale analysis

In order to understand, if genotype dependent growth in repeat spawning fish (relative to fist time spawners) occurs prior to or after the first spawning event, using a subset of individuals that were assigned to focal populations ($N_{repeat.spawners}$=200, $N_{first.time.spawners}$=378, 78.1 and 75.3 % of total assigned individuals, respectively), we compared scale growth of fish at corresponding subsections of the scale (see also Figure S1) Scale growth in salmon have high correlation to growth and have been used as a surrogate measure to monitor growth (ICES, 2011), as well as to monitor growth at a specific time period (e.g. (Aykanat *et al.*, 2015). Scale growth is highly correlated to actual fish growth, both in the juvenile and the adult phases in Atlantic salmon (i.e. r>0.90 for juvenile and adults from the Teno River, e.g. (Aykanat *et al.*, 2015; Erkinaro *et al.*, 2018). In our dataset, the total scale length within repeat spawners were also highly correlated to length at capture (r=0.83), and with a similar slope to that of the first time spawning counterparts (Figure S1). Similarly, seasonal growth of fish was used to infer growth of fish at a specific period in their lifetime, measured according to standard guidelines provided by ICES (2011). We analyzed scale growth before and after the first spawning event until the fish was capture during the second spawning event, as the proxy for growth in the same period (See also Figure S1).